BOUNDED CONJUGATORS FOR REAL HYPERBOLIC AND UNIPOTENT ELEMENTS IN SEMISIMPLE LIE GROUPS

ANDREW SALE

ABSTRACT. Let G be a real semisimple Lie group with trivial centre and no compact factors. Given a conjugate pair of either real hyperbolic elements or unipotent elements a and b in G we find a conjugating element $g \in G$ such that $d_G(1,g) \leq L(d_G(1,u) + d_G(1,v))$, where L is a positive constant which will depend on some property of a and b. For the vast majority of such elements however, L can be assumed to be a uniform constant.

CONTENTS

The objective of this paper is to present results concerning an effective version of the conjugacy problem in the setting of semisimple Lie groups. We focus on finding short conjugators between real hyperbolic elements and unipotent elements.

The conjugacy problem is one of Max Dehn's three decision problems in group theory, which he set out in 1912, motivated by questions in low-dimensional manifolds. The other problems are the word problem and the isomorphism problem. These three problems are fundamental in realms of combinatorial and geometric group theory and have received much attention over the last century. Dehn originally described these problems in group theory because of the significance he discovered they had in the geometry of 3–manifolds. He observed the interplay that occurs between the fundamental group of the manifold and its geometry. For example, the conjugacy problem in the fundamental group is equivalent to determining when two loops in the manifold are freely homotopic.

Let Γ be a recursively presented group with finite symmetric generating set A. The *word problem* on Γ asks whether there is an algorithm which determines when any given word on the generating set A represents the identity element of Γ. Associated to the word problem is the Dehn function, which measures its geometric complexity. It is a measure of the minimal area required to fill a loop in the Cayley 2–complex of G. Because of this geometric interpretation, determining the Dehn function of groups has been a fundamental question in geometric group theory over

the last couple of decades. The extra information the Dehn function provides means that we could describe calculating it as an *effective* version of the word problem.

The *conjugacy problem* is of a similar flavour to the word problem. We say the conjugacy problem in Γ is solvable if there is an algorithm which, on input two words u and v on the generating set A, determines whether u and v represent conjugate elements in Γ.

An effective version of the conjugacy problem. Estimating the length of short conjugating elements in a group could be described as an *effective* version of the conjugacy problem. Suppose a group G admits a left-invariant metric d_G. For $g \in G$ let $|g|$ denote $d_G(1, g)$. The *conjugacy length function* is the minimal function $\mathrm{CLF}_G : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ which satisfies the following: for $x \in \mathbb{R}_{\geq 0}$, if u is conjugate to v in G and $|u| + |v| \leq x$ then there exists a conjugator $g \in G$ such that $|g| \leq \mathrm{CLF}_G(x)$. One can define it more concretely to be the function which sends $x \in \mathbb{R}_{\geq 0}$ to

$$\sup \big\{ \inf\{|g| : gu = vg\} : |u| + |v| \leq x \text{ and } u \text{ is conjugate to } v \text{ in } G \big\}.$$

The question of determining conjugacy length functions has been addressed previously. See for example [Sal12c] and [Sal12a] for results concerning groups including free solvable groups, wreath products, group extensions and abelian-by-cyclic groups. We know that the conjugacy length function is linear for groups including hyperbolic groups [BH99], Right-angled Artin groups [CGW09] and Mapping class groups [MM00], [BD11], [Tao11]. For CAT(0)–groups and biautomatic groups all we know is that it is at most exponential (see [BH99]), it is an open question as to whether this bound is sharp and indeed we do not even know if it is not necessarily linear. In their work on the stronger ℓ^1–Bass conjecture, Ji, Ogle and Ramsey show 2–step nilpotent groups have a quadratic conjugacy length function [JOR10], and also obtain a result for relatively hyperbolic groups. The fundamental group of a prime 3–manifold also has a quadratic upper bound [BD11], [Sal12a].

The reader should note that, unlike the conjugacy problem itself, in order to define the conjugacy length function the only requirement on Γ is that it should admit a left-invariant metric. In particular this means that we can define it for Lie groups.

In this paper we take the first steps towards understanding the conjugacy length function of higher-rank real semisimple Lie groups and their lattices by studying the conjugacy of real hyperbolic elements and unipotent elements.

Grunewald and Segal solved the conjugacy problem in arithmetic groups [GS80] and hence, by Margulis arithmeticity, we know that every lattice in a higher-rank real semisimple Lie group has solvable conjugacy problem. However, as discussed in [GI05], their solution gives no insight into the length of the conjugating element. Determining a control on the lengths of short conjugators in lattices would not only be interesting in its own right, but would also provide us with a method of proving the solubility of the conjugacy problem via a method which does not rely on the power of Margulis arithmeticity.

Real hyperbolic elements. Analogies have frequently been drawn between lattices in higher-rank semisimple Lie groups and mapping class groups. The pseudo-Anosov elements of a mapping class group are the elements which behave in a similar way to the real hyperbolic elements. Recently J. Tao [Tao11] showed that mapping class groups have linear conjugacy length functions, however earlier work of Masur and Minsky [MM00] showed that there is a linear bound on the length of short conjugators between a pair of pseudo-Anosov elements. Hence, following the analogy through to lattices, it is natural to begin approaching this problem by studying the real hyperbolic elements. The analogy also carries through to

$\mathrm{Out}(F_n)$, the outer automorphism group of a free group F_n. Here we do not yet know if the conjugacy problem is solvable, however Lustig [Lus07] has shown that it is solvable when we restrict to iwip elements, which are the elements analogous to the pseudo-Anosov and hyperbolic elements.

Theorem 1 below describes the main result of this paper for real hyperbolic elements. Let G be a real semisimple Lie group with trivial centre and no compact factors. An element $a \in G$ is said to be real hyperbolic if it translates some geodesic in the associated symmetric space X and furthermore it also translates every other geodesic parallel to the first. The slope of a describes how these geodesics sit inside the Weyl chambers of X. We formalise these definitions in Section 1.

Theorem 1. *For each slope ξ there exist positive constants ℓ_ξ and d_ξ such that if a and b are conjugate real hyperbolic elements in G with slope ξ and such that $d_X(p, ap), d_X(p, bp) \geq d_\xi$ then there exists a conjugator $g \in G$ satisfying:*

$$d_X(p, gp) \leq 2\ell_\xi\big(d_X(p, ap) + d_X(p, bp)\big).$$

This leads naturally to the following result for lattices:

Theorem 2. *Let Γ be a lattice in G. Then for each slope ξ, there exists a constant L_ξ such that two elements $a, b \in \Gamma$ are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq L_\xi\big(d_X(p, ap) + d_X(p, bp)\big).$$

The main tool in proving Theorem 1 concerns an estimate of the distance from an arbitrary basepoint in X to a flat (or union of flats):

Lemma 3.6 *For each slope ξ there exist positive constants ℓ_ξ and d_ξ such that for each $a \in G$ which is real hyperbolic of slope ξ and such that $d_X(p, ap) > d_\xi$ the following holds:*

$$d_X(p, \mathrm{MIN}(a)) \leq 2\ell_\xi d_X(p, ap).$$

Unipotent elements. In the second half of this paper we present a method for dealing with the conjugacy of certain unipotent elements. Crucial to our method is the root-space decomposition of the Lie algebra $\mathfrak{g}$ of G. This gives us a root system Λ such that a maximal unipotent subgroup N will have a Lie algbera $\mathfrak{n}$ of the form

$$\mathfrak{n} = \sum_{\lambda \in \Lambda^+} \mathfrak{g}_\lambda$$

where $\mathfrak{g}_\lambda$ is a root-space of $\mathfrak{g}$ and Λ^+ is a subset of positive roots in Λ. In the following we assume that $\mathfrak{g}$ is a split Lie algebra, meaning that each root-space $\mathfrak{g}_\lambda$ has dimension 1. The classical split Lie algebras are $\mathfrak{sl}_d(\mathbb{R})$, $\mathfrak{sl}_{d,d+1}(\mathbb{R})$, $\mathfrak{sp}_d(\mathbb{R})$ and $\mathfrak{so}_{d,d}(\mathbb{R})$. Because the exponential map restricted to $\mathfrak{n}$ is a diffeomorphism, each element u of N can be expressed uniquely as

$$u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right), \quad Y_\lambda \in \mathfrak{g}_\lambda.$$

If $\Pi \subset \Lambda^+$ is the set of simple roots, then we say that the *simple entries* of u are those Y_λ for $\lambda \in \Pi$.

Theorem 3. *Fix $\delta > 0$. Let u and v be conjugate unipotent elements in G such that each simple entry of u and v is of size at least δ. Then there exists $g \in G$ such that $gug^{-1} = v$ and which satisfies:*

$$d_G(1, g) \leq L(d_G(1, u) + d_G(1, v))$$

where L depends on δ and the root-system Λ associated to G.

As with the real hyperbolic case we are able to deduce a corollary for lattices:

Theorem 4. *Let Γ be a lattice in G. Then there exists a constant $L > 0$ such that two unipotent elements u and v in Γ with non-zero simple entries are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$d_G(1, g) \leq L\big(d_G(1, u) + d_G(1, v)\big).$$

From here, there are several natural directions one could look. Firstly one could try to remove the restriction that the Lie algebra should be split. Secondly, it seems that the techniques to prove Theorem 3 could be extended to a larger family of unipotent elements in G, allowing the simple entries to be zero. In particular this would potentially allow the result for lattices to include all unipotent elements. However the computational aspect involved in doing this increased many times over from the "simple case" considered in this paper.

The next step should be to consider elliptic elements and then the idea would be to use the (complete) Jordan decomposition, which expresses each element of the group as a product of commuting elliptic, real hyperbolic and unipotent elements.

Question: Given the Jordan decompositions of two conjugate elements u and v of G, what can we infer about the length of a short conjugator between u and v from the lengths of short conjugators between each of the Jordan components?

The centralisers of each component will play an important role in answering this question.

The nature of Theorems 2 and 4 raise further questions for lattices. In particular we would like the conjugating element to come from the lattice, rather than the Lie group as we have here.

Question: Consider two elements u and v in Γ which are conjugate in the lattice. Suppose we have a conjugator g for u and v such that g lies in the ambient Lie group. How close to g is a conjugating element from Γ?

We will now outline the structure of this paper. In Section 1 we discuss the relevant background information related in particular to the structure of symmetric spaces. Sections 2-5 deal with real hyperbolic elements, while Sections 6-8 tackle the problem for unipotent elements.

In Section 2 we describe the geometry of the centralisers of real hyperbolic elements of G. Theorems 1 and 2 are the main objectives of Section 3. In Section 4 we consider the question of whether there is a uniform linear bound for the length of short conjugators between all real hyperbolic elements in G. We answer the question in the negative. Section 5 discusses the issue of translating the conjugator found in Theorem 2, which lies in the ambient Lie group G, into the lattice. In specific cases we show this can be done without losing the linear control on conjugator length. However the question of what happens in general, even for real hyperbolic elements, remains open.

Section 6 introduces the problem for unipotent elements, and in particular includes a brief overview of what happens when we apply our method to upper-triangular unipotent matrices. In Section 7 we show how the root system of G plays a vital role in the conjugacy of unipotent elements. Finally, construction of the short conjugator is done in Section 8.

1. Preliminaries

Let G be a semisimple real Lie group and let X be its associated symmetric space. For background on the structure of symmetric spaces we refer the reader to [Hel01] and [Ebe96]. The author's thesis [Sal12b] contains the contents of this paper, and as such the preliminaries described there should also provide the required background material.

1.1. **Symmetric spaces and their isometries.** Given an isometry g of a symmetric space X (or any CAT(0)–space) we can consider the following set

$$\mathrm{MIN}(g) = \left\{ x \in X \mid d_X(x, gx) = \inf_{q \in X} d_X(q, gq) \right\}.$$

If $\mathrm{MIN}(g)$ is non-empty then we say that g is a *semisimple* isometry of X, otherwise it is called parabolic. The semisimple isometries which fix some point in X are called *elliptic*. Various names have been given to those which don't fix a point, including hyperbolic, axial and loxodromic. To minimise confusion, we will simply call these elements the *non-elliptic semisimple* elements of G.

It is not hard to see that the non-elliptic semisimple isometries of X will translate some geodesic (this is shown, for example, in [BGS85]). Given a bi-infinite geodesic $c : \mathbb{R} \to X$ consider $P(c)$, the subspace of X consisting of all geodesics that are parallel to c. We call an isometry g of X *real hyperbolic* if g translates some geodesic c in X and furthermore $\mathrm{MIN}(g) = P(c)$. This says precisely that any geodesic parallel to c will also be translated by g. The nomenclature used for these elements is justified in [Sal12b, §3.1.3 & Lemma 3.2.3].

Let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition of the Lie algebra $\mathfrak{g}$ of G. There exists a point $p \in X$ such that the subalgebra $\mathfrak{k}$ is the Lie algebra of the maximal compact subgroup $K = G_p = \{g \in g \mid gp = p\}$. Meanwhile $\mathfrak{p}$ can be identified with T_pX, the tangent space at p of X. A geodesic $c : \mathbb{R} \to X$ such that $c(0) = p$ determines a vector in T_pX and hence an element $H \in \mathfrak{p}$. Consider the maximal abelian subspace $\mathfrak{a}$ of $\mathfrak{p}$ which contains H. The submanifold $\exp(\mathfrak{a})p$ is a maximal flat in X. This follows from the following two facts:

- [Hel01, Ch. IV Thm 7.2] for $\mathfrak{s} \subset \mathfrak{p}$, the submanifold $\exp(\mathfrak{s})p$ in X is totally geodesic in X if and only if $\mathfrak{s}$ is a Lie triple system;
- [Hel01, Ch. IV Thm 4.2] for $Y_1, Y_2, Y_3 \in \mathfrak{p}$, the curvature tensor at p is given by $R_p(Y_1, Y_2)Y_3 = -[[Y_1, Y_2], Y_3]$.

Lemma 1.1. *Every maximal flat F containing p is of the form $F = \exp(\mathfrak{a})p$ for some maximal abelian subspace $\mathfrak{a}$ of $\mathfrak{p}$.*

Proof. This follows from [Hel01, Ch. V Prop 6.1]. $\qquad\square$

1.1.1. *Weyl chambers.* In order to define the Weyl chambers of a flat in X we first need to describe the *root-space decomposition* of the Lie algebra of G. Fix a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Let $\mathfrak{a}$ be a maximal abelian subspace of $\mathfrak{p}$. Then, since the operators $\mathrm{ad}(H)$, for $h \in \mathfrak{a}$, are simultaneously diagonalisable we can consider, for linear functionals $\lambda : \mathfrak{a} \to \mathbb{R}$, the eigenspaces

$$\mathfrak{g}_\lambda = \{Y \in \mathfrak{g} \mid \mathrm{ad}(H)Y = \lambda(H)Y \text{ for all } H \in \mathfrak{a}\}.$$

Those λ for which $\mathfrak{g}_\lambda$ is non-empty are called *roots* of $\mathfrak{g}$ with respect to $\mathfrak{a}$ and the spaces $\mathfrak{g}_\lambda$ the corresponding *root-spaces*. Let Λ be the set of all non-zero roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$. The *root-space decomposition* of $\mathfrak{g}$ is the following:

$$\mathfrak{g} = \mathfrak{g}_0 + \sum_{\lambda \in \Lambda} \mathfrak{g}_\lambda.$$

The roots Λ form a root system in the dual space $\mathfrak{a}^*$. A subset Π of Λ is called a *base* if it is a basis for $\mathfrak{a}^*$ and if any root λ can be written as

$$\lambda = \sum_{\alpha \in \Pi} c_\alpha \alpha$$

in such a way that either each c_α is non-negative or each c_α is non-positive. The elements of Π are called *simple roots* and those elements for which $c_\alpha \geq 0$ for each $\alpha \in \Pi$ are called *positive roots* with respect to Π. We denote the set of positive roots by Λ^+.

Consider a flat F in X. By Lemma 1.1 there exists a maximal abelian subspace $\mathfrak{a}$ of $\mathfrak{p}$ such that $F = \exp(\mathfrak{a})p$. Let Λ be the corresponding set of roots, Π a set of simple roots and Λ^+ the corresponding positive roots. The set of elements $H \in \mathfrak{a}$ for which $\lambda(H) > 0$ for each $\lambda \in \Pi$ forms an open *Weyl chamber* in $\mathfrak{a}$, denoted $\mathfrak{a}^+$. The corresponding set $\mathcal{C}_\Pi = \exp(\mathfrak{a}^+)p$ is called an open Weyl chamber in X. The choice of Π determines the Weyl chamber $\mathfrak{a}^+$.

For each root $\lambda \in \Lambda$ the kernel is a hyperplane in $\mathfrak{a}$. These are called the *singular hyperplanes* of $\mathfrak{a}$. The *walls* of $\mathcal{C}_\Pi$ are contained in the singular hyperplanes and are defined, for a subset $\Theta \subset \Pi$, as

$$\mathcal{C}_\Theta = \{\exp(H) \in \overline{\mathcal{C}_\Pi} \mid \lambda(H) = 0 \text{ for } H \in \Pi \setminus \Theta\}$$

where $\overline{\mathcal{C}_\Pi}$ is the closure of $\mathcal{C}_\Pi$ in F. The flat F is partitioned into Weyl chambers and walls. In fact, after removing all the singular hyperplanes from F, the connected components of what remains are all the Weyl chambers corresponding to the different choices for Π.

Let $c : \mathbb{R} \to X$ be a geodesic in F with $c(0) = p$. Then there exists $H \in \mathfrak{a}$ such that $c(t) = \exp(tH)$ for every $t \in \mathbb{R}$. If H is contained in a singular hyperplane of $\mathfrak{a}$ then we say the geodesic c is *singular*. Otherwise H is contained in some Weyl chamber $\mathcal{C}_\Pi$ and we call c a *regular geodesic* in X. Since every geodesic in X is contained in some maximal flat this definition extends to all geodesics. The following is an equivalent definition of regular and singular geodesics:

Proposition 1.2. *A geodesic is regular if and only if it is contained in a unique maximal flat.*

Proof. See [Ebe96, §2.11]. □

1.1.2. *Slopes of geodesics.* Given two geodesic rays ρ_1, ρ_2 in X, we say they are *asymptotic* if they are at finite Hausdorff distance from one-another. This defines an equivalence relation on geodesic rays in X, the equivalence classes of which form the *ideal boundary* $\partial_\infty X$ of X. The action of an isometry $g \in G$ on X can be extended to an action on $\partial_\infty X$ since ρ_1 and ρ_2 are asymptotic if and only if $g\rho_1$ and $g\rho_2$ are asymptotic. Hence we may consider the quotient of the action of G on $\partial_\infty X$. We denote this quotient by Δ_{mod}. The Δ_{mod}–*direction*, or *slope*, of a ray ρ is the image of ρ under the quotient maps.

Consider a bi-infinite geodesic $c : \mathbb{R} \to X$. This determines two boundary points, one for each end of the geodesic. Although, by the definition of a symmetric space, there is an isometry $\varphi = s_{c(0)}$ of X, the geodesic involution at $c(0)$, such that $\varphi c(t) = c(-t)$ for all $t \geq 0$, this isometry will not be in the connected component of the group of isometries of X, and thus not in G. Thus the two ends of c determine two ideal points corresponding to $c(\infty)$ and $c(-\infty)$, and these will usually give rise to distinct slopes. We call the *slope of the bi-infinite geodesic $c : \mathbb{R} \to X$* the projection of $\sigma(\infty)$ onto Δ_{mod}.

If the geodesic c is regular, then we say the corresponding Δ_{mod}–directions are regular, while if c is singular its slopes are said to be singular too. Equivalently,

the regular Δ_{mod}–slopes are the ones contained in the interior of Δ_{mod}, while the singular slopes are those in the boundary of Δ_{mod}.

1.1.3. *Families of parallel geodesics.* We defined in Section 1.1.2 the slope of a geodesic. Given a geodesic $c : \mathbb{R} \to X$, its slope ξ belongs to the set Δ_{mod}, which is the closure of a model chamber in the boundary of X. If the slope of c is contained in the interior of Δ_{mod} then it is regular and therefore contained in a unique maximal flat F. In fact, by the Flat Strip Theorem [BH99, Pg. 182], any geodesic parallel to c will also be contained in F. Hence F is equal to the subspace of X containing all geodesics parallel to c.

If c is a singular geodesic in X then it will be contained in a whole family of maximal flats. Again, using the Flat Strip Theorem, any geodesic parallel to c must be contained in one of these flats.

More formally, let $P(c)$ denote the subspace of X consisting of all geodesics that are parallel to c. So when c is regular, $P(c)$ is equal to the unique maximal flat F described above. While if c is singular $P(c)$ will be the union of (infinitely many) maximal flats. The structure of these sets is discussed in more detail in [Ebe96, §2.20].

Lemma 1.3. *G acts transitively on the set of subspaces of the form $P(\sigma)$, where σ varies over geodesics with the same slope.*

Proof. Let $\sigma, \tau : \mathbb{R} \to X$ be non-parallel geodesics of the same slope and let F_1, F_2 be any pair of maximal flats containing σ, τ respectively. The transitivity of the action of G on the set of maximal flats in X is well known and follows from [Hel01, Ch. V Thm 6.4]. Further more we know there exists $g \in G$ such that $gF_1 = F_2$ and $g\sigma(0) = \tau(0)$. We now have two geodesics, $g\sigma$ and τ, which are contained in the same maximal flat and have the same slope. If the positive rays, that is $g\sigma[0, \infty)$ and $\tau[0, \infty)$, are in the same Weyl chamber, then having the same slope implies the geodesics must coincide, hence $g\sigma = \tau$. If they are not in the same Weyl chamber then we apply an element of the Weyl group of F_2, which acts transitively on the Weyl chambers, so that they end up in the same Weyl chamber. An element of the Weyl group is a coset of the point-wise stabiliser of F_2. So by choosing a representative of this coset we have $k \in K = G_{\tau(0)}$ such that $kg\sigma = \tau$. Finally, if σ' is any geodesic parallel to σ, then $kg\sigma'$ will be parallel to $kg\sigma = \tau$. Hence $kgP(\sigma) \subseteq P(\tau)$ and equality follows by symmetry. □

1.2. **Asymptotic cones.** We briefly discuss here asymptotic cones and a result of Kleiner and Leeb about the asymptotic cones of a symmetric space. Before we define an asymptotic cone we should discuss ultralimits and ultrafilters.

A *non-principal ultrafilter* ω on $\mathbb{N}$ is a finitely additive probability measure on $\mathbb{N}$ which takes values of either 0 or 1 and all finite sets have zero measure. Given a sequence $(a_n)_{n \in \mathbb{N}}$ of real numbers the *ultralimit* of this sequence is $a = \lim_\omega (a_n) \in \mathbb{R}$ which has the property that $\omega\{n \in \mathbb{N} \mid |a_n - a| < \varepsilon\} = 1$ for every $\varepsilon > 0$.

Let X be a metric space with metric d and let $p = (p_n)_{n \in \mathbb{N}}$ be a sequence of points in X. Let $(d_n)_{n \in \mathbb{N}}$ be a sequence in $(0, \infty)$ which diverges to infinity. Given a non-principal ultrafilter ω on $\mathbb{N}$ we can define the *asymptotic cone* $\mathrm{Cone}_\omega(X, p, d_n)$ to be the quotient space of sequences $(x_n)_{n \in \mathbb{N}}$ such that $\lim_\omega \frac{d(p_n, x_n)}{d_n} < \infty$ under the equivalence relation saying that two sequences (x_n) and (y_n) are equivalent if and only if $\lim_\omega \frac{d(x_n, y_n)}{d_n} = 0$. We can define a metric d_ω on the cone by setting $d_\omega((x_n), (y_n)) = \lim_\omega \frac{d(x_n, y_n)}{d_n}$. A point $x \in \mathrm{Cone}_\omega(X, p, d_n)$ is said to be the ultralimit of a sequence of points (x_n) in X if (x_n) is a member of the equivalence class determining x.

As explained in, for example, [BGS85, Appendix 5], the ideal boundary of a symmetric space can be given a spherical building structure. If we take an asymptotic cone of a symmetric space X then, as Kleiner and Leeb put it in [KL97], intuitively speaking we are "pulling the spherical building structure from infinity to the space of directions." Spherical buildings have a useful property, which can be described as *rigidity of angles*. This means that given two points in a spherical building, their location inside their respective chambers determines a finite set of possible (angular) distances between them. When this property is pulled to the tangent space, by taking the asymptotic cone of our symmetric space, it transfers to a similar statement regarding the angle between intersecting geodesics.

We refer the reader to [KL97] for the definition of a Euclidean building used by Kleiner and Leeb. However, note that in her PhD thesis [Par00] Parreau showed that Kleiner and Leeb's axioms are equivalent to the definition of a building given by Tits.

Theorem 1.4 (Kleiner–Leeb [KL97]). *Let X be a symmetric space of noncompact type. For any sequence of positive numbers (d_n) diverging to infinity and for any sequence of points $p = (p_n)$ in X the asymptotic cone $\mathrm{Cone}_\omega(X, p, d_n)$ is a Euclidean building modelled on the Euclidean Coxeter complex (E, M), where E is $\mathrm{rank}(X)$–dimensional Euclidean space and M is the quotient of the set-wise stabiliser $\mathrm{Stab}_G(E)$ by the point-wise stabiliser $\mathrm{Fix}_G(E)$.*

1.3. **A note on lattices.** A *lattice* in a semisimple real Lie group G is a discrete subgroup Γ such that $\Gamma \backslash G$ has finite volume with respect to the Haar measure on G. If the quotient $\Gamma \backslash G$ is compact, then we say Γ is a *cocompact* or *uniform* lattice in G. Otherwise we say it is *non-uniform*. A lattice is said to be *irreducible* if, whenever G is given as a product of Lie groups $G_1 \times G_2$, then the projections of Γ into each factor are dense.

Suppose that Γ is irreducible. Let d_Γ denote a word metric on Γ with respect to some finite generating set. We can also consider the size of an element of Γ using a Riemannian metric d_G on G. By a theorem of Lubotzky, Mozes and Raghunathan [LMR00], provided the real rank of G is at least 2, it does not matter which we use. Furthermore, this means that if we fix a basepoint p in the symmetric space X then we could also use $d_\Gamma(p, \gamma p)$ to estimate $d_\Gamma(1, \gamma)$.

2. Centralisers of real hyperbolic elements

Let g be a non-elliptic semisimple isometry of X and suppose $\sigma : \mathbb{R} \to X$ is a geodesic translated by g, oriented so that $g\sigma(0) = \sigma(t)$ for some $t > 0$. We define the *slope* of g to the be Δ_{mod}–direction of σ corresponding to the positive direction, $\sigma(\infty)$. Of course parallel geodesics are asymptotic and so we get the same Δ_{mod}–direction regardless of which geodesic we consider. We say an isometry g is *regular semisimple* if it is non-elliptic semisimple and its slope is regular in Δ_{mod}. When g is non-elliptic semisimple but its slope is singular we say g is a *singular semisimple* isometry.

Lemma 2.1. *Let a be a regular semisimple element in G, contained in a maximal torus A. Then:*
  (1) there exists a unique maximal flat F_a in X which is stabilised by a;
  (2) for any $p \in F_a$, let K be the stabiliser of p, then $\mathrm{Stab}(F_a) = Z_G(a)N_K(A)$, where $Z_G(a)$ is the centraliser of a in G and $N_K(A)$ is the normaliser of A in K.

Proof. Suppose a is regular semisimple, translating a geodesic c. We can decompose a as $a = hk = kh$ where h is real hyperbolic and k is elliptic. The real hyperbolic

component will translate c, hence $\mathrm{MIN}(h)$ is equal to the maximal flat F_a. Consider the action of $k = h^{-1}a$ on F_a. It will fix c pointwise and hence must act trivially on F_a since any non-trivial elliptic isometry on F_a will permute the Weyl chambers and therefore cannot fix a regular geodesic. But this tells us that the action of a on F_a is precisely the same as that of h, thus $\mathrm{MIN}(a) = F_a$.

Suppose that another flat F' is stabilised by a. Since a is not elliptic , it must act on F' hyperbolically, by translating some geodesic in F'. This geodesic must therefore be contained in $\mathrm{MIN}(a) = F$. So F and F' intersect in a regular geodesic, hence must be equal. This proves (1).

Let p, K be as in (2) and note that $F_a = Ap$. Let $z \in Z_G(a)$ and $w \in N_K(A)$. Then $zwF_a = zF_a$ and $azF_a = zaF_a = zF_a$, so by (1) $zF_a = F_a$. Hence $Z_G(a)N_K(A) \subseteq \mathrm{Stab}(F_a)$. Now let $g \in \mathrm{Stab}(F_a)$. Then there exists $b \in A$ such that $gp = bp$. Thus $b^{-1}g = k \in K$ and stabilises F_a. Since $kAk^{-1}p = F_a$ we see that kAk^{-1} is a maximal torus stabilising the flat F_a. Such a torus is unique, so $k \in N_K(A)$. Hence $g = bk \in Z_G(a)N_K(A)$ and (2) holds. $\square$

We can see that the orbit of the centraliser of a regular semisimple element will be a maximal flat. Consider instead a singular semisimple element that is also real hyperbolic element. We may call such elements *singular real hyperbolic*. The following Lemma tells us that the orbit of the centraliser of a singular real hyperbolic element will contain many maximal flats.

Lemma 2.2. *Let a be a singular real hyperbolic element in G. Then:*

   (1) the subspace $\mathrm{MIN}(a)$ is precisely the set of all geodesics translated by a, which is the Riemannian product of a Euclidean space and a symmetric space of noncompact type; and
   (2) for any $p \in \mathrm{MIN}(a)$, let K be the stabiliser of p, then
$$Z_G(a) \subseteq \mathrm{Stab}(\mathrm{MIN}(a)) \subseteq Z_G(a)K$$
   where $Z_G(a)$ is the centraliser of a in G.

Proof. The first part of assertion (1) follows from the definition we gave for a real hyperbolic element, while for a proof of the latter part of (1) we refer the reader to [Ebe96, 2.11.4].

For (2), take $b \in Z_G(a)$ and let c' be any geodesic translated by a. Then $bc' = bac' = abc'$ implies that bc' is translated by a, hence is contained in $\mathrm{MIN}(a)$. Now let g stabilise $\mathrm{MIN}(a)$. Let c_1, c_2 be the geodesics translated by a such that $c_1(0) = p$ and $c_2(0) = gp$ respectively. Since c_1 and c_2 are parallel they are contained in a common flat $F = Ap$, where A is a maximal abelian Lie subgroup of G which contains a. There exists $b \in A$ such that $bc_2 = c_1$ and $bgp = p$. Hence $bg \in K$ and in particular $g \in Z_G(a)K$. $\square$

3. Finding a short conjugator for real hyperbolic elements

The aim of this section is to obtain a control on the length of a conjugator between two real hyperbolic elements a, b in G. The control will be linear, but the constant in the upper bound will depend on the slope of a and b, and hence their conjugacy class. Our method of demonstrating this is to first show that a conjugator corresponds to an isometry that maps $\mathrm{MIN}(a)$ to $\mathrm{MIN}(b)$. Then, by obtaining a control on the distance from an arbitrary basepoint p to $\mathrm{MIN}(a)$ in terms of $d_X(p, ap)$, we can obtain a control on the length of a conjugator from G.

3.1. **Relating conjugators to maps between flats.** Here we show why we can obtain a short conjugator by understanding the distance to $\mathrm{MIN}(a)$ and $\mathrm{MIN}(b)$. In the following let π_a be the orthogonal projection of X onto $\mathrm{MIN}(a)$ and for $x \in X$ let $G_q = \{k \in G \mid gq = q\}$.

Proposition 3.1. *Let a, b be conjugate non-elliptic semisimple elements in G. Then:*

  (1) for $g \in G$, if $gag^{-1} = b$ then $g\mathrm{MIN}(a) = \mathrm{MIN}(b)$;
  (2) for $h \in G$, if $h\mathrm{MIN}(a) = \mathrm{MIN}(b)$ then there exists $x \in G_{\pi_a(p)}$ such that $(hx)a(hx)^{-1} = b$.

Proof. For (1), let c be any geodesic stabilised by a. Since $g = bga^{-1}$ we see that the geodesic gc is translated by b and so is contained in $\mathrm{MIN}(b)$. Hence $g\mathrm{MIN}(a) \subseteq \mathrm{MIN}(b)$. Similarly we get $g^{-1}\mathrm{MIN}(b) \subseteq \mathrm{MIN}(a)$ and (1) is proved.

Next suppose that $g, h \in G$ are such that $gag^{-1} = b$ and $h\mathrm{MIN}(a) = \mathrm{MIN}(b)$. By the first part we observe $g^{-1}h\mathrm{MIN}(a) = \mathrm{MIN}(a)$, so $g^{-1}h \in \mathrm{Stab}(\mathrm{MIN}(a)) \subseteq Z_G(a)K$, with the latter relationship coming from Lemma 2.2, where we take $K = G_{\pi_a(p)}$. Then there exists $x \in K$ such that $g^{-1}hx \in Z_G(a)$. This implies $(g^{-1}hx)a(g^{-1}hx)^{-1} = a$ and so $(hx)a(hx)^{-1} = gag^{-1} = b$, proving (2). □

If a and b are regular semisimple elements in G then they each stabilise a unique maximal flat in X. Suppose a, b stabilise maximal flats F_a and F_b respectively. By taking our basepoint p to be in F_a we can build a quadrilateral which has two vertices in F_a and two vertices in F_b, as in Figure 1. In light of 3.1, the aim is to find an element $g \in G$ of a controlled size which maps the flat F_a to F_b. We will then obtain a conjugator of controlled size.
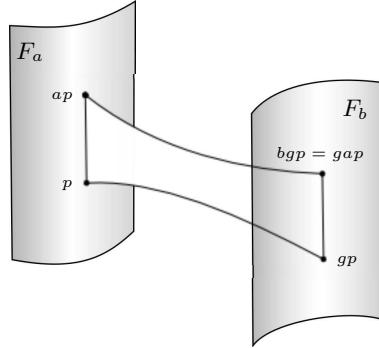
FIGURE 1. A quadrilateral in X demonstrating the conjugacy of a and b in G.

Proposition 3.2. *Let p be our basepoint in X. Suppose for all semisimple a in G we can find a constant $\ell(a)$ such that:*

$$d_X(p, \mathrm{MIN}(a)) \leq \ell(a)d_X(p, ap).$$

Then for a, b conjugate hyperbolic elements in G there exists a conjugator $g \in G$ such that:

$$d_X(p, gp) \leq \ell(a)d_X(p, ap) + \ell(b)(p, bp).$$

Proof. Choose points $p_a \in \mathrm{MIN}(a)$ and $p_b \in \mathrm{MIN}(b)$ which satisfy

$$d_X(p, p_a) \leq \ell(a)d_X(p, ap) \quad \text{and} \quad d_X(p, p_b) \leq \ell(b)d_X(p, bp).$$

Let $g_1 \in G$ be such that $g_1\mathrm{MIN}(a) = \mathrm{MIN}(b)$ and $g_1p_a = p_b$. By Lemma 3.1 there exists x in G_{p_b} such that $(xg_1)a(xg_1)^{-1} = b$. Let $g = xg_1$.
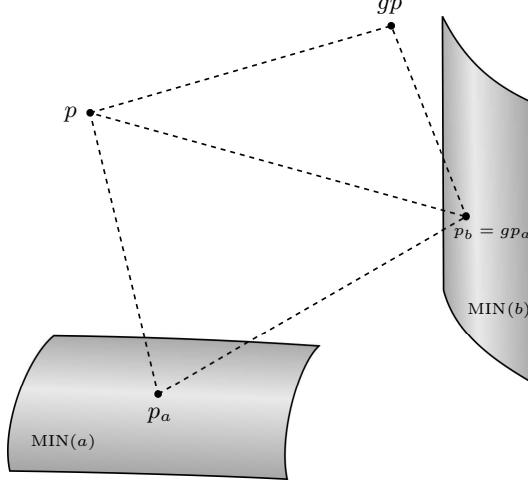
FIGURE 2. Obtaining an upper bound on $d_X(p, gp)$.

To finish the proof we need to check that we have the required upper bound on $d_X(p, gp)$. By the triangle inequality:

$$
\begin{aligned}
d_X(p, gp) &\leq d_X(p, p_b) + d_X(p_b, gp) \\
&= d_X(p, p_b) + d_X(gp_a, gp) \\
&= d_X(p, p_b) + d_X(p, p_a)) \\
&\leq \ell(a)d_X(p, ap) + \ell(b)d_X(p, bp).
\end{aligned}
$$

$\square$

3.2. **Bounding the distance to a flat.** In Proposition 3.2 we saw how finding some constant $\ell(a)$ such that $d_X(p, \mathrm{MIN}(a)) \leq \ell(a)d_X(p, ap)$ helps us to control the size of a conjugator in G between a and another element b. The aim of this section is to find such constants for the case when a is real hyperbolic. In fact, we will determine a value for $\ell(a)$ which depends only on the slope of a. Note that we have the following:

Lemma 3.3. *If a is conjugate to b then a and b have the same slope.*

Proof. Let a have slope $\xi \in \Delta_{\mathrm{mod}}$. This means that the geodesic segment $[p, ap]$, where p is a point in $\mathrm{MIN}(a)$, has Δ_{mod}-direction ξ. Suppose $b = gag^{-1}$ for some $g \in G$. Then g maps the bi-infinite geodesic through p and ap to a bi-infinite geodesic through gp and $gap = bgp$. This geodesic is translated by b, so the slope of b is the Δ_{mod}-direction of the geodesic segment $[gp, bgp] = g[p, ap]$. Since the Δ_{mod}–direction is defined to be G–invariant we have that the slope of b is ξ. $\square$

In order to determine the value of $\ell(a)$ we will use an asymptotic cone of X, which, by a result of Kleiner and Leeb [KL97], is a Euclidean building. It is helpful therefore to first determine the corresponding value in a Euclidean building. This will then be useful to find the value for symmetric spaces. Recall that a real hyperbolic element a satisfies $\mathrm{MIN}(a) = P(\sigma)$, where σ is any geodesic translated by a and $P(\sigma)$ is the subspace of X consisting of all geodesics parallel to σ. We are therefore interested in the distance to similarly defined subspaces of a Euclidean building. Also recall that, for a Euclidean building Y, the quotient of $\partial_\infty Y$ by the group of isometries of Y is denoted by Δ_{mod}, and $\theta : \partial_\infty Y \to \Delta_{\mathrm{mod}}$ is the natural map.

Lemma 3.4 (see [HKM10, Lemma 5.1])**.** *Let Y be a Euclidean building, $\delta \in \partial_\infty Y$, c be a geodesic in Y with one end asymptotic to δ and E be the subset of Y consisting of all geodesics parallel to c. Then for any point $p \in Y$ the geodesic ray emanating from p which is asymptotic to δ enters the set E in finite distance.*

Remark: The proof of this Lemma given in [HKM10, Lemma 5.1] only covers algebraic Euclidean buildings. It is worth noting that by [KT04] the asymptotic cone of a symmetric space is an algebraic Euclidean building, so their proof applies to the buildings which we are concerned with.

Proposition 3.5. *Let a be an isometry of a Euclidean building Y such that a translates all geodesics parallel to a geodesic c and let E be the subset of Y containing all these geodesics. Suppose the ray $c[0, \infty)$ satisfies $ac[0, \infty) \subset c[0, \infty)$ and is asymptotic to $\delta \in \partial_\infty E$, where $\theta(\delta) = \xi \in \Delta_{\mathrm{mod}}$. Then there exists a constant ℓ_ξ such that for any basepoint p in Y the following holds:*

$$d(p, E) \leq \ell_\xi d(p, ap).$$
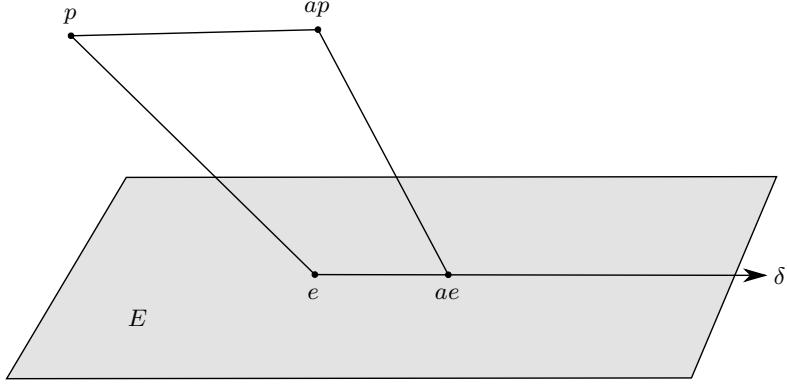
FIGURE 3. The angle $\angle_{ae}(e, ap)$ is rigid, that is the angle is contained in a finite set which is determined by ξ. This leads to a bound on the distance from p to E.

Proof. Consider the ray emanating from p which represents δ. By Lemma 3.4 this ray enters E. Let e be the first point along this ray such that $e \in E$. By design ae also lies on this ray. Now translate the ray by a. What we get is a geodesic triangle in Y, as seen in Figure 3, with vertices p, ae, ap. By the rigidity of angles in Y, $\angle_{ae}(p, ap)$ belongs to the finite set $D(\xi)$. Let ϕ be minimal in this set. Then since Y is a CAT(0) space we have that $d(p, ap) \geq d(ap, ae) \sin \phi$. Hence we put $\ell_\xi = \frac{1}{\sin \phi}$ and the proposition holds. $\square$

Lemma 3.6. *Let the Tits building structure on $\partial_\infty X$ have anisotropy polyhedron Δ_{mod}. Fix a basepoint p in X. Then for each element $\xi \in \Delta_{mod}$ there exists positive constants ℓ_ξ and d_ξ such that for each $a \in G$ which is real hyperbolic of slope ξ and such that $d_X(p, ap) > d_\xi$ the following holds:*

$$d_X(p, \mathrm{MIN}(a)) \leq 2\ell_\xi d_X(p, ap).$$

Proof [regular slopes]. The proof which follows applies to the case when ξ is regular. The proof for singular slopes is analogous, with slight modifications which are described at the end of this proof.

First note that the constant ℓ_ξ will be the same constant that we obtained in Proposition 3.5.

We proceed by contradiction, supposing the statement is false. We then obtain a sequence of regular semisimple elements a_n in G, each of slope ξ, such that $d_X(p, a_n p)$ diverges to infinity and which satisfies:

$$d_X(p, F_n) > 2\ell_\xi d_X(p, a_n p) \tag{1}$$

where F_n is the unique maximal flat stabilised by a_n. Write $d_n := d_X(p, a_n p)$ and $D_n := d_X(p, F_n)$. Let $\pi_n : X \to F_n$ be the orthogonal projection onto F_n and let t_n be the translation length of a_n, that is $t_n := d_X(\pi_n(p), a_n \pi_n(p))$. We split the proof into three parts depending on the limits of the ratios t_n/D_n and d_n/D_n.

CASE 1: $\lim_\omega(t_n/D_n) \neq 0 \neq \lim_\omega(d_n/D_n)$.

In the first part of the proof we build a Euclidean building and use Proposition 3.5 to obtain a contradiction under the assumption that t_n/D_n does not converge to zero.

Since d_n diverges to infinity, it follows from (1) that D_n does too. Hence we pick a non-principal ultrafilter ω and consider the asymptotic cone $Y = \mathrm{Cone}_\omega(X, p, D_n)$. By choice of scalars it follows that the ultralimit E of the sequence of flats F_n is contained in Y and lies a distance 1 away from the point p (when we view p as an element of the cone Y).

Define the map $g : Y \to Y$ by sending $(x_n) \in Y$ to $(a_n x_n)$. To check it is well-defined on Y we need only observe that it moves p a bounded distance:

$$
\begin{aligned}
d_\omega(p, gp) &= \lim_\omega\left(\frac{d_X(p, a_n p)}{D_n}\right) \\
&= \lim_\omega\left(\frac{d_n}{D_n}\right) \\
&\leq \lim_\omega\left(\frac{1}{2\ell_\xi}\right) \\
&= \frac{1}{2\ell_\xi}.
\end{aligned}
$$

Furthermore, since a_n acts on X by isometries for each n it follows that g acts on Y by isometries.

By assumption t_n/D_n does not converge to zero, hence $d_\omega(\pi(p), g\pi(p)) > 0$. It implies, since $t_n \leq d_n$, that $d_X(p, gp) > 0$. Under these conditions we may apply Proposition 3.5, since g acts on E by translating along geodesics towards a boundary point ξ. This gives us the following contradiction:

$$1 = d_\omega(p, E) \leq \ell_\xi d_\omega(p, gp) \leq \frac{\ell_\xi}{2\ell_\xi} = \frac{1}{2}.$$

CASE 2: $\lim_\omega(t_n/D_n) = 0 = \lim_\omega(d_n/D_n)$.

We must therefore assume the ω-limit of t_n/D_n is zero. We assume this for the second part of the proof and we also assume that the ω-limit of d_n/D_n is zero. We will build a sequence of quadrilaterals and take their Hausdorff limit. The limiting quadrilateral will be flat and intersecting a flat F only in one edge, along a regular geodesic. This will give the contradiction.

Fix a flat F in X and a point $q \in F$. For each n consider an isometry $g_n \in G$ which sends $\pi_n(p)$ to q and F_n to F. The first thing to note is that each geodesic segment $[\pi_n(p), a_n \pi_n(p)]$ is mapped to a geodesic segment $T := [q, g_n a_n g_n^{-1} q]$ in F of Δ_{mod}-direction ξ. Consider the projection $\tau : [g_n p, g_n a_n p] \to T$. For a fixed constant h and for large enough n we may pick a subsegment S of T of length h such that the pre-image $\tau^{-1}(S)$ in $[g_n p, g_n a_n p]$ has length at most $d_n h/t_n$.

Label points $b_n^{(1)}, b_n^{(2)}$ on the geodesic $[g_np, a_ng_np]$ which are mapped under τ to each end of the segment S, see Figure 4. Let $L_n^{(i)} := d_X(b_n^{(i)}, T)$ and without loss of generality assume $L_n^{(1)} \leq L_n^{(2)}$. Observe:

$$D_n = d_X(g_np, T) \leq d_X(g_np, b_n^{(1)}) + d_X(b_n^{(1)}, T)$$

and replacing g_np by g_na_np if necessary we get:

$$L_n^{(1)} \geq D_n - \frac{d_n}{2}.$$

Using our hypothesis we therefore get $L_n^{(1)} > \left(2\ell_\xi - \frac{1}{2}\right) d_n$. So, referring back to the value of ℓ_ξ obtained in Proposition 3.5, since $2\ell_\xi - \frac{1}{2} > 0$ for any choice of ξ, we get that $L_n^{(1)}$ diverges to infinity.

Now define quadrilaterals Q_n as follows. We take one edge to be the segment S and the two adjacent edges are those subsegments of $[b_n^{(i)}, \tau(b_n^{(i)})]$ of length h which include the points $\tau(b_n^{(i)})$, for $i = 1, 2$. The quadrilateral Q_n has three sides of length h. Let $\psi_n(h)$ be the length of the fourth side. Define a map $\hat{\psi}_n : [0, L_n^{(1)}] \to \mathbb{R}$ measuring the distance across the flat rhombus (see Figure 5).

Using the CAT(0) property of X we see that

$$
\begin{aligned}
h \leq \psi_n(h) \quad &\leq \quad \hat{\psi}_n(h) \\
&= \quad \left(\frac{L_n^{(1)} - h}{L_n^{(1)}}\right) h + \frac{h^2 d_n}{L_n^{(1)} t_n} \\
&= \quad h - \frac{h^2}{L_n^{(1)}} + h^2 \left(\frac{d_n}{D_n t_n}\right)\left(\frac{D_n}{L_n^{(1)}}\right)
\end{aligned}
$$
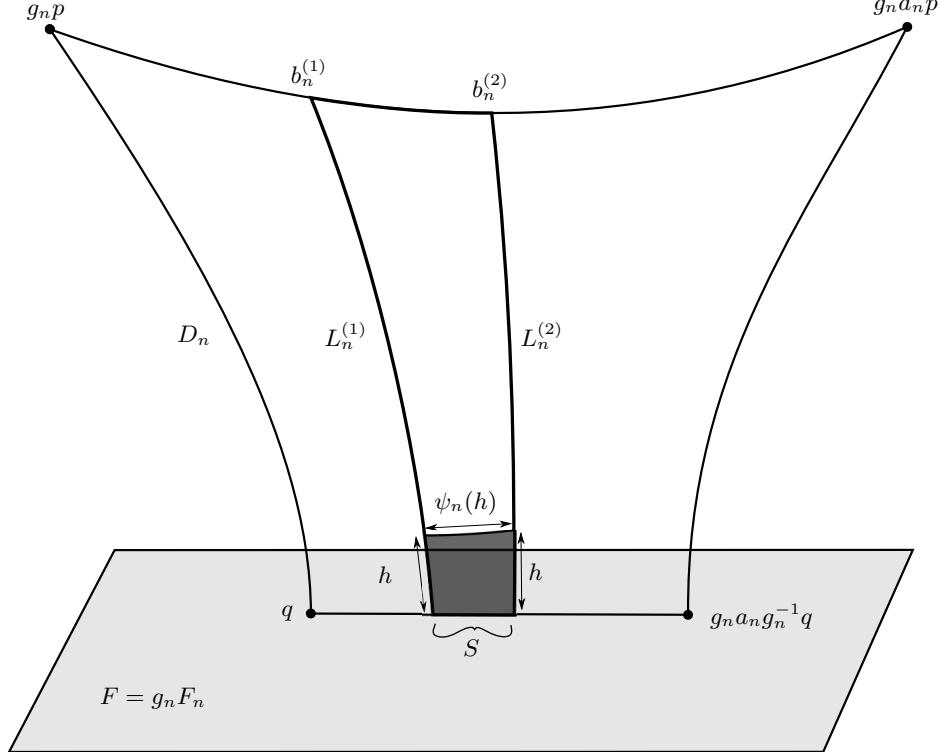
FIGURE 4. The quadrilateral Q_n with side S.

Above we showed that $L_n^{(1)} \geq D_n - \frac{d_n}{2}$. We can use this to show that $D_n/L_n^{(1)}$ is bounded above by $\left(1 - \frac{1}{4\ell_\xi}\right)^{-1}$. This is therefore enough, since we have the assumption that d_n/D_n converges to zero, to conclude that $\psi_n(h)$ converges to h as n tends to infinity. After translating the quadrilaterals Q_n along the geodesics to q we can find the Hausdorff limit Q of a convergent subsequence of the quadrilaterals Q_n. The quadrilateral Q will have four sides with length h, two right-angles and hence must be a flat square. But Q intersects F only through the regular geodesic segment T. This gives a contradiction.

<u>CASE 3:</u> $\lim_\omega(t_n/D_n) = 0 \neq \lim_\omega(d_n/D_n)$.

We conclude by combining both of the above arguments into one in order to obtain a contradiction when t_n/D_n converges to zero but d_n/D_n does not. We start by looking at the situation inside the Euclidean building Y that we built in case 1. It is constructed so that $d_\omega(p, E) = 1$, but in what follows we will show that in this case we would have the contradiction $d_\omega(p, E) < 1$.

In Y, take the two geodesic rays asymptotic to ξ which begin at p and at gp respectively. Note that the second ray is the image of the first under g. Also recall that both rays will enter the apartment E. Since g fixes E pointwise we see that the two rays must come together at some point y. In particular either y is the point where the rays enter E or it is not in E. We will show that $d_\omega(y, E) \leq (4\ell_\xi)^{-1}$.

Suppose that $d_\omega(y, E) > (4\ell_\xi)^{-1}$ and let (y_n) be a sequence of points in X which represent y in Y. Then there exists ε such that $(4\ell_\xi)^{-1} < \varepsilon < d_\omega(y, E)$ and $\omega\{n \in \mathbb{N} \mid d_X(y_n, F_n) \geq \varepsilon D_n\} = 1$. We now proceed as in case 2, for each n applying the isometry g_n and constructing quadrilaterals Q_n with one edge in the fixed flat F. As before we pick a segment S of length h from $\left[q, g_n a_n g_n^{-1} q\right]$ whose pre-image under the projection onto $\left[q, g_n a_n g_n^{-1} q\right]$ intersects $[g_n y_n, g_n a_n y_n]$ in a segment of length at least $\frac{h}{t_n} d_X(y_n, a_n y_n)$. Let $L_n^{(i)}$ be the distances between the corresponding end-points of these subsegments and suppose $L_n^{(1)} \leq L_n^{(2)}$. Then in order to proceed as before we need that the function:

$$\hat{\psi}_n(t) = \left(\frac{L_n^{(1)} - t}{L_n^{(1)}}\right) h + \frac{t}{L_n^{(1)}} \frac{h}{t_n} d_X(y_n, a_n y_n)$$

converges to h when we put $t = h$. We need to check two things: firstly that $L_n^{(1)}$ diverges to infinity and secondly that $\frac{d_X(y_n, a_n y_n)}{L_n^{(1)}}$ converges to zero. For the former
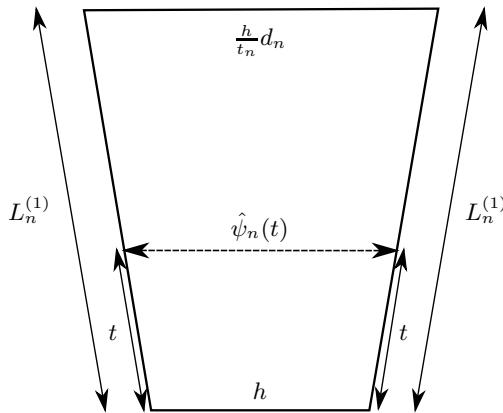

FIGURE 5. Defining $\hat{\psi}_n(t)$.
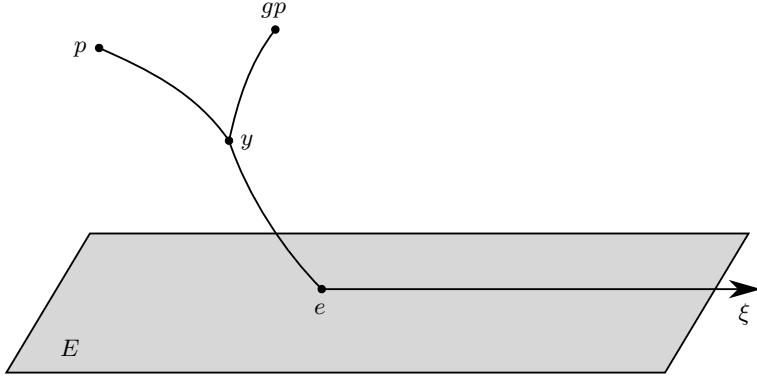
FIGURE 6. Two geodesic rays asymptotic to ξ, entering the apartment E at a point e and merging at the point y.

we note that for all but finitely many $n \in \mathbb{N}$ we have the following:

$$
\begin{aligned}
L_n^{(1)} &\geq d_X(y_n, F_n) - \frac{1}{2} d_X(y_n, a_n y_n) \\
&> \left(2\ell_\xi \varepsilon - \frac{1}{2} \right) d_X(y_n, a_n y_n)
\end{aligned}
$$

If $d_X(y_n, a_n y_n)$ is bounded we use the first line to show $L_n^{(1)}$ is unbounded. Otherwise we use the second line, recalling that $\varepsilon > \frac{1}{4\ell_\xi}$. To prove that $\frac{d_X(y_n, a_n y_n)}{L_n^{(1)}}$ converges to zero we first check that $\frac{D_n}{L_n^{(1)}}$ is bounded. This is so because for all but finitely many n we have the following:

$$
\begin{aligned}
L_n^{(1)} &\geq d_X(y_n, F_n) - \frac{1}{2} d_X(y_n, a_n y_n) \\
&> \varepsilon D_n - \frac{1}{2} d_n \\
\frac{L_n^{(1)}}{D_n} &> \varepsilon - \frac{d_n}{2 D_n} \\
\frac{D_n}{L_n^{(1)}} &< \left(\varepsilon - \frac{1}{4\ell_\xi} \right)^{-1}
\end{aligned}
$$

Hence we see that $\frac{d_X(y_n, a_n y_n)}{L_n^{(1)}} = \frac{d_X(y_n, a_n y_n)}{D_n} \frac{D_n}{L_n^{(1)}}$ converges to zero by our choice of y_n. Then as before, after translating the quadrilaterals Q_n so they each have a vertex at q, we take the Hausdorff limit of a convergent subsequence of these quadrilaterals and obtain a flat quadrilateral which intersects F only through a regular geodesic segment. Here we have our contradiction and conclude that $d_\omega(y, E) \leq (4\ell_\xi)^{-1}$.

To finish the argument we look at the triangle in Y with vertices p, gp, y. In a similar manner to the proof of Proposition 3.5 we use the fact that Y is a CAT(0) space to get $d_\omega(p, gp) \geq d_\omega(p, y)\ell_\xi^{-1}$, recalling that ℓ_ξ^{-1} is the sine of the minimal angle in the finite set $D(\xi)$ of possible angles between geodesics of Δ_{mod}-direction ξ.

Therefore we have the following:

$$
\begin{aligned}
d_\omega(p, E) &\leq d_\omega(p, y) + d_\omega(y, E) \\
&\leq \ell_\xi d_\omega(p, gp) + \frac{1}{4\ell_\xi} \\
&\leq \ell_\xi \frac{1}{2\ell_\xi} + \frac{1}{4\ell_\xi} \\
&= \frac{1}{2} + \frac{1}{4\ell_\xi} < 1
\end{aligned}
$$

However, we also have $d_\omega(p, E) = \lim_\omega \left(\frac{D_n}{D_n} \right) = 1$, thus giving the contradiction and proving the Lemma. $\qquad\square$

Proof [singular slopes]. In order to modify the above proof to work for singular directions we need to make the following adjustments. First we replace the flats F_n by $\mathrm{MIN}(a_n)$. Case 1 continues as above with no change. For cases 2 and 3, the contradiction we obtain will be similar. Instead of using a fixed flat F, we use a fixed subspace $P(c)$ which consists of a family of geodesics parallel to some geodesic c of slope ξ, for example we may take $M = \mathrm{MIN}(a_1)$ and c any geodesic translated by a_1. By Lemma 1.3 we know there exists $g_n \in G$ which sends $\mathrm{MIN}(a_n)$ to M. From the proof of Lemma 1.3 it is also clear that g_n can be chosen so it sends a geodesic translated by a_n to a geodesic translated by a_1. Furthermore, if we fix a point q in M, as we did in the above proof, then we can choose g_n so it sends $\pi_n(p)$ to q. Once we have this, we can find a flat quadrilateral Q in the same way as above, but it will intersect M only in one side, which is a geodesic segment of slope ξ. The opposite edge of Q will be a segment of a geodesic parallel to c, hence should be contained in M, but it is not. $\qquad\square$

In light of Lemma 3.6 we can put $\ell(a) = \ell(b) = 2\ell_\xi$, where ξ is the slope of a and b, into Proposition 3.2 to get the following:

Theorem 3.7. *Let ℓ_ξ and d_ξ be the constants from Lemma 3.6. Suppose a and b are conjugate real hyperbolic elements in G with slope $\xi \in \Delta_{\mathrm{mod}}$ and such that $d_X(p, ap), d_X(p, bp) \geq d_\xi$. Then there exists a conjugator $g \in G$ such that:*

$$
d_X(p, gp) \leq 2\ell_\xi \big(d_X(p, ap) + d_X(p, bp) \big).
$$

The constant $2\ell_\xi$ that we have obtained will depend on the slope of a and b, and hence on the conjugacy class. However it is important to note that it is independent of the basepoint p that was chosen.

When we restrict our attention to a lattice Γ in G, Theorem 3.7 will apply to all but finitely many real hyperbolic elements of Γ. This leads to the following result:

Corollary 3.8. *Let Γ be a lattice in G. Then for each $\xi \in \Delta_{\mathrm{mod}}$, there exists a constant L_ξ such that two elements $a, b \in \Gamma$ are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$
d_X(p, gp) \leq L_\xi \big(d_X(p, ap) + d_X(p, bp) \big).
$$

We now offer a couple of corollaries to Theorem 3.7 which shed a little more light on the nature of short conjugators between real hyperbolic elements in a semisimple real Lie group. For a semisimple element a of G, the translation length of a is

$$
\tau(a) = \inf\{ d_X(x, ax) \mid x \in X \}.
$$

We can reformulate Theorem 3.7 so that it applies to all real hyperbolic elements in G, provided their translation length isn't too small.

Corollary 3.9. *Let $0 < \varepsilon \leq d_\xi$ and suppose that a and b are real hyperbolic elements of G of slope ξ and with translation lengths $\tau(a), \tau(b) \geq \varepsilon$. Then a and b are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq 2\ell_\xi \left(\frac{d_\xi}{\varepsilon} + 1 \right) \big(d_X(p, ap) + d_X(q, aq) \big).$$

Proof. Since a is real hyperbolic, $\tau(a) > 0$ and $\tau(a^k) = k\tau(a)$ for all $k \in \mathbb{N}$. Let k to be the maximal positive integer such that $k\varepsilon < d_\xi$. In particular, maximality of k implies that $d_X(p, a^{k+1}p), d_X(p, b^{k+1}p) \geq d_\xi$ and we are able to apply Lemma 3.6, concluding that

$$d_X(p, \mathrm{MIN}(a)) \leq 2\ell_\xi(k+1)d_X(p, ap), \quad d_X(p, \mathrm{MIN}(b)) \leq 2\ell_\xi(k+1)d_X(p, bp).$$

We can then apply Lemma 3.2, taking $\ell(a) = \ell(b) = 2\ell_\xi(\frac{d_\xi}{\varepsilon} + 1)$. This gives the upper bound on the length of a conjugator g as required. $\qquad\square$

We complete this section with the following consequence of the above work. It says that if we restrict ourselves to looking at the majority of regular semisimple elements — that is, those with not too small translation length and of a slope which is not too close to being singular — then we can obtain a linear bound on the length of short conjugators.

Corollary 3.10. *For every $\varepsilon_1, \varepsilon_2 > 0$ there exists $\kappa = \kappa(\varepsilon_1, \varepsilon_2)$ with the following property: assume that a and b are conjugate hyperbolic elements with translation lengths $\tau(a), \tau(b) \geq \varepsilon_1$ and slope $\xi \in \Delta_{\mathrm{mod}}$ such that the spherical distance from ξ to $\partial\Delta_{\mathrm{mod}}$ is at least ε_2. Then there exists a conjugator $g \in G$ such that:*

$$d_X(p, gp) \leq \kappa\big(d_X(p, ap) + d_X(p, bp)\big).$$

It is worth noting that while we have a precise expression for the constant ℓ_ξ in terms of the slope ξ, we have no grasp on the value taken by d_ξ.

4. Dependence on the slope

The aim here is to show that any constant satisfying the linear relationship in Theorem 3.7 must depend on the common slope of a and b. To do this, we first show there is not a uniform constant $\ell > 0$ such that for all regular semisimple elements $a \in G$ the following holds:

$$(2) \qquad\qquad d_X(p, \mathrm{MIN}(a)) \leq \ell d_X(p, ap)$$

where p is an arbitrary basepoint in X. This will imply that the constants $\ell(a)$ required for Proposition 3.2 will have to depend somehow on a.

To do this, we will construct a sequence of regular semisimple elements which contradict the existence of such an ℓ. The sequence will converge to a singular semisimple element, agreeing with the intuition that the constant ℓ_ξ of Lemma 3.6 diverges to infinity as the slope ξ converges to a singular direction.

We first note that if (2) is true for a point $p \in X$ then it is true for every point $q \in X$. Indeed let $g \in G$ be any isometry such that $gp = q$. Then firstly:

$$
\begin{aligned}
d_X(q, \mathrm{MIN}(a)) &= d_X(gp, \mathrm{MIN}(a)) \\
&= d_X(p, g^{-1}\mathrm{MIN}(a)) \\
&= d_X(p, \mathrm{MIN}(g^{-1}ag))
\end{aligned}
$$

and secondly:

$$
\begin{aligned}
d_X(q, aq) &= d_X(gp, agp) \\
&= d_X(p, g^{-1}agp).
\end{aligned}
$$

But since we assume (2) to be true for all hyperbolic elements it follows that it is true for $g^{-1}ag$ and thus:

$$d_X(q, \mathrm{MIN}(a)) \leq \ell d_X(q, aq).$$

Fix a pair of distinct flats F, F' whose intersection is non-trivial and of dimension at least one. Let p_0 be a point in their intersection and let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be the corresponding Cartan decomposition. Let $\mathfrak{a}, \mathfrak{a}'$ be the maximal abelian subspaces of $\mathfrak{p}$ such that:

$$F = \exp(\mathfrak{a})p_0$$
$$F' = \exp(\mathfrak{a}')p_0$$

Furthermore suppose $H \in \mathfrak{a} \cap \mathfrak{a}'$ has length 1 and $Y \in \mathfrak{a}' \setminus \mathfrak{a}$ is orthogonal to H and is also of length 1.

Let (H_n) be a sequence of regular unit vectors in $\mathfrak{a} \setminus \mathfrak{a}'$ which converge to H. Define $a_0 := \exp(H) \in G$ and $a_n := \exp(H_n) \in G$ for $n \in \mathbb{N}$ (see Figure 7).
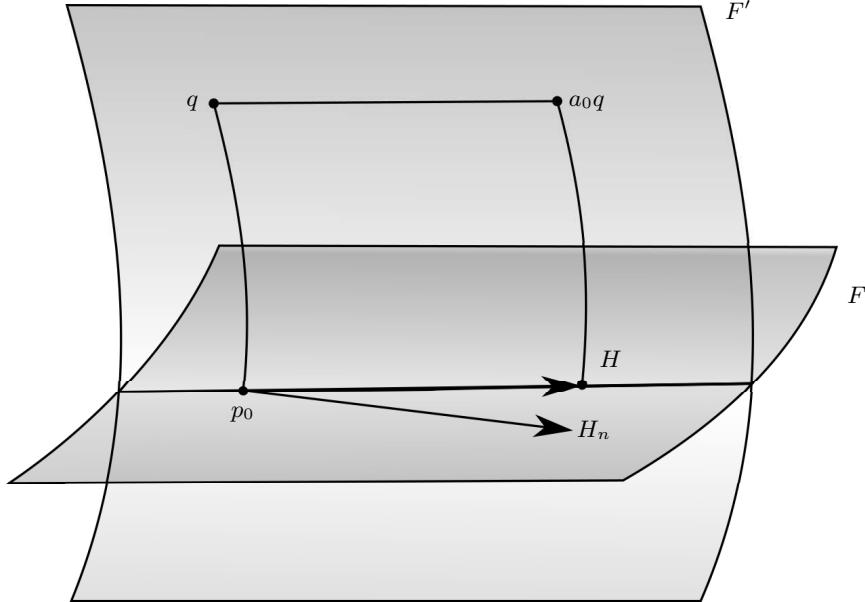
FIGURE 7. The two flats F and F'.

We suppose (2) holds for some $\ell > 0$. Let φ be the angle between Y and the flat F and set $q = \exp(\frac{2\ell}{\sin \varphi} Y)p_0$. Then

$$d_X(q, F) \geq \sin \varphi \, d_X(q, p_0) = 2\ell.$$

But for each $n \in \mathbb{N}$, by construction, $F = \mathrm{MIN}(a_n)$. Thus $d_X(q, \mathrm{MIN}(a_n)) \geq 2\ell$ for each n. Meanwhile the sequence of points $a_n q$ converges to the point $a_0 q$ as n tends to infinity. Hence $d_X(q, a_n q)$ converges to 1. This gives the following contradiction:

$$
\begin{aligned}
2\ell &\leq & d_X(q, \mathrm{MIN}(a_n)) \\
&\leq & \ell d_X(q, a_n q) \\
&\to & \ell
\end{aligned}
$$

Hence there cannot exist such a constant ℓ which satisfies (2) for every hyperbolic element in G.

The following Lemma explains how to use the above to demonstrate the non-existence of a uniform constant for the linear control on conjugacy length among all regular semisimple elements in G.

Lemma 4.1. *Suppose for $p \in X$ there exists $\ell' > 0$ such that for every pair of conjugate regular semisimple elements a, b in G there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq \ell'(d_X(p, ap) + d_X(p, bp)).$$

Then (2) holds for $\ell = 2\ell'$.

Proof. Let a be regular semisimple in G such that $\mathrm{MIN}(a) = F_a$ does not contain p. Let π_a be the orthogonal projection of X onto F_a and let m be the midpoint of the geodesic segment $[p, \pi_a(p)]$. Consider the geodesic symmetry s_m of X about m, that is the map $s_m : X \to X$ such that for any geodesic $c : \mathbb{R} \to X$ with $c(0) = m$, $s_m(c(t)) = c(-t)$ for all $t \in \mathbb{R}$. Since X is a symmetric space s_m is an isometry. Let F denote the flat $s_m(F_a)$. Then $p \in F$ and $[p, \pi_a(p)]$ meets both flats at right-angles, so this geodesic segment realises the distance between the flats.

Take $g \in G$ such that $gF = F_a$ and $gp = \pi_a(p)$. Then by the hypothesis of the Lemma, $d_X(p, \pi_a(p)) \leq \ell'(d_X(p, ap) + d_X(p, g^{-1}agp))$. But $d_X(p, g^{-1}agp)$ is the translation length of a, so is less than $d_X(p, ap)$. Hence $d_X(p, F_a) \leq 2\ell' d_X(p, ap)$. $\qquad\square$

Since (2) cannot hold, we have the following:

Corollary 4.2. *For $p \in X$ there does not exists $\ell' > 0$ such that for every pair of conjugate regular semisimple elements a, b in G there exists a conjugator $g \in G$ such that*

$$d_X(p, gp) \leq \ell'(d_X(p, ap) + d_X(p, bp)).$$

5. Looking for a short conjugator in Γ

In Corollary 3.8 we found a short conjugator between two real hyperbolic elements in Γ. However this conjugator lies in the ambient Lie group G. In order to improve our understanding of conjugacy length in Γ we need to work out how to move our conjugator from G so that it becomes a conjugator in Γ. The main obstacle here is in understanding how the lattice will intersect flats in the symmetric space.

Given a conjugator $g \in G$ for $a, b \in \Gamma$, the set of all conjugators is the coset $Z_G(a)g$ of the centraliser of a. We are therefore interested in the contents of the set $Z_G(a)g \cap \Gamma$, or equivalently $gZ_G(b) \cap \Gamma$. If we begin with the assumption that a conjugator for a, b from the lattice exists, then at least we know these sets are non-empty.

When a and b are real hyperbolic we have a good understanding of the geometry of their centralisers (see Lemmas 2.1 and 2.2). For example, when b is regular we can find a point $q \in X$ such that the orbit $Z_G(b)q$ is a maximal flat. Then $gZ_G(b)q$ is also a maximal flat and is in fact the unique maximal flat stabilised by a. Hence it is equal to $Z_G(a)gq$. Clearly $(Z_G(a)g \cap \Gamma)q$ is contained in this flat, and the question is how far is gq from this subset? Once we know this distance, we can shift our conjugator g, whose length we have an estimate for courtesy of Section 3, to a conjugator in Γ and keep track of the size of the new lattice conjugator.

Suppose $\gamma \in gZ_G(b) \cap \Gamma$. Then $gZ_G(b) \cap \Gamma = \gamma(Z_G(b) \cap \Gamma)$. It is therefore enough to look at how the set $(Z_G(b) \cap \Gamma)q$ sits inside $Z_G(b)q$. In the singular case $Z_G(b)q$ will be made up of a family of maximal flats. Each flat will be the orbit of a maximal torus T contained in $Z_G(b)$. If there exists some such torus T which satisfies the properties that $b \in T \cap \Gamma$ and $T \cap \Gamma$ is isomorphic to $\mathbb{Z}$, then we understand what the fundamental domain for the action of $T \cap \Gamma$ on Tq will look like: it will be an R–tubular neighbourhood of a hyperplane orthogonal to the geodesics translated by b, where $2R \leq d_X(q, bq)$.

This situation cannot arise if the $\mathbb{Q}$–rank of the lattice is too small: it must satisfy $\mathrm{rank}_{\mathbb{Q}}(\Gamma) \geq \mathrm{rank}_{\mathbb{R}}(G) - 1$. If $\mathrm{rank}_{\mathbb{Q}}(\Gamma) = \mathrm{rank}_{\mathbb{R}}(G) - 1$ then a maximal $\mathbb{Q}$–split torus S is a hyperplane inside a maximal $\mathbb{R}$–split torus T. Because Γ must intersect S in a finite set, Γ will intersect T in nothing more than a finite extension of $\mathbb{Z}$. In particular, if $q \in X$ is chosen so that Tq is flat, then $(\Gamma \cap T)q$ will look like a copy of $\mathbb{Z}$ inside the flat. Furthermore, the fundamental domain for the action of $T \cap \Gamma$ on this flat will be a tubular neighbourhood of Sq.

In general, if $\mathrm{rank}_{\mathbb{Q}}(\Gamma) = \mathrm{rank}_{\mathbb{R}}(G) - d$, then the flats in X which have a non-trivial intersection with an orbit of Γ can do so only with copies of $\mathbb{Z}^k$ for $\mathrm{rank}_{\mathbb{R}}(G) \geq k \geq d$.

Lemma 5.1. *Suppose* $\mathrm{rank}_{\mathbb{Q}}(\Gamma) \geq \mathrm{rank}_{\mathbb{R}}(G) - 1$ *and let T be a maximal $\mathbb{R}$–split torus in G such that $T \cap \Gamma$ is a finite extension of $\mathbb{Z}$. Take a real hyperbolic element $b \in T \cap \Gamma$ and suppose it is conjugate in Γ to a. If a and b have slope ξ, then there exists a conjugator $\gamma \in \Gamma$ for a, b such that*

$$d_X(p, \gamma p) \leq (6L_\xi + 1)(d_X(p, ap) + d_X(p, bp))$$

where L_ξ is as in Corollary 3.8.
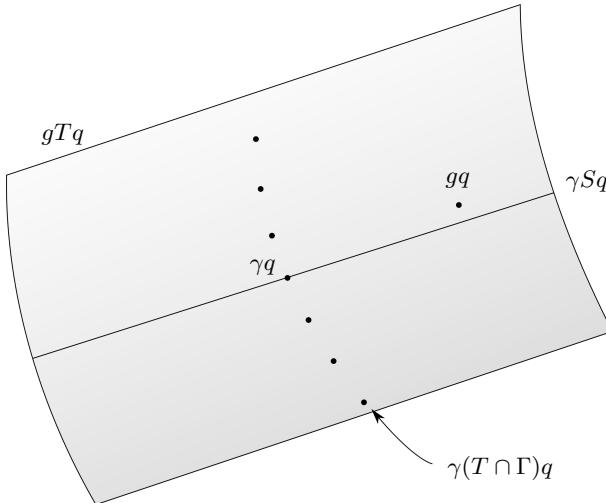
FIGURE 8. When $\mathrm{rank}_{\mathbb{R}}(G) = 2$, the quotient of the maximal flat gTq by the action of $\Gamma \cap gTg^{-1}$ is a cylinder whose diameter is bounded above by $d_X(\gamma q, b\gamma q) = d_X(q, aq)$.

Proof. Suppose first that $\mathrm{rank}_{\mathbb{Q}}(\Gamma) = \mathrm{rank}_{\mathbb{R}}(G) - 1$ and let S be a maximal $\mathbb{Q}$–split torus contained in T. To adapt the following to the case when $\mathrm{rank}_{\mathbb{Q}}(\Gamma) = \mathrm{rank}_{\mathbb{R}}(G)$, we merely take S^+ to instead be the face of a Weyl chamber in T (which will be isometric to a $\mathbb{Q}$–Weyl chamber, so S^+ will still isometrically embed into $\Gamma \backslash X$).

Let q be any point in X such that Tq is a (maximal) flat. Note that Tq will be contained in $\mathrm{MIN}(b)$. Choose $g \in G$ as in Theorem 3.7, taking $p = q$. Then g maps Tq to a maximal flat contained in $\mathrm{MIN}(a)$ and in particular there exists $\gamma \in \Gamma$ such that gq lies in an R–tubular neighbourhood of γSq, where $2R = d_X(q, aq)$. In particular there is a $\mathbb{Q}$–Weyl chamber S^+ in S such that gq lies in an R–tubular neighbourhood of $\gamma S^+ q$.

The $\mathbb{Q}$–Weyl chamber maps isometrically into $\Gamma \backslash X$ (see [Leu04]), so in particular $d_X(\gamma q, gq) \leq d_X(\Gamma q, gq) + R$. Hence $d_X(q, \gamma q) \leq 2d_X(q, gq) + R$.

To finish, we need to translate it so it works for an arbitrary basepoint p. To do this we use Lemma 3.6 and apply the triangle inequality:

$$
\begin{aligned}
d_X(p, \gamma p) &\leq 2d_X(p, q) + d_X(q, \gamma q) \\
&\leq 4L\xi d_X(p, ap) + 4L_\xi(d_X(q, aq) + d_X(q, bq)) + r \\
&\leq (6L_\xi + 1)(d_X(p, ap) + d_X(p, bp))
\end{aligned}
$$

Note that in the above we have assumed that $d_X(p, ap) \leq d_X(p, bp)$; if this is not the case, we can just reverse the roles of a and b. $\qquad\square$

The preceding Lemma works because we know enough about the shape and size of a fundamental domain for the action of $\Gamma \cap T$ on the flat Tq. In this case, Γq intersects Tq in a copy of $\mathbb{Z}$, and we essentially have a tubular neighbourhood of a co-dimension 1 flat, the radius of the neighbourhood bounded above by the translation length of a.

Question: When Γq intersects Tq in a copy of $\mathbb{Z}^k$, for some $k \geq 2$, what can we say about the dimensions of the fundamental domain for the action of $\Gamma \cap T$ on the k–dimensional flat inside Tq stabilised by $\Gamma \cap T$ in terms of $d_X(p, bp)$ and $d_X(p, \mathrm{MIN}(b))$, where $b \in T \cap \Gamma$?

6. Unipotent elements

We move on now to look at the conjugacy of unipotent elements in G. The result obtained here gives a linear bound on the length of a short conjugator from G between two unipotent elements which satisfy a certain algebraic condition. Provided the two elements are, in some sense, pushed far enough away from certain root-spaces in $\mathfrak{g}$, the linear bound obtained will be uniform. The method used to prove it relies heavily on the Lie algebra and in particular on the root system corresponding to $\mathfrak{g}$.

Consider two conjugate elements u, v in N. When looking at the case when N is the subgroup of $\mathrm{SL}_n(\mathbb{R})$ consisting of the unipotent upper triangular matrices, the condition we impose on u and v is equivalent to demanding that the super-diagonal entries in the matrix, that is the $(i, i+1)$–entries, are all not too close to zero. The short conjugator we obtain is built up by gradually knocking off entries in the matrix until you are left with a matrix with zeros above the super-diagonal. Doing this for both u and v gives two matrices that are then related via conjugation by a diagonal matrix. In Section 6.1 we give a few more details about how the process works for $\mathrm{SL}_n(\mathbb{R})$.

In the process used to knock off the extra terms in the matrix, the super-diagonal entries play an important role and it is crucial that they are not too close to zero. The underlying root system for $\mathrm{SL}_n(\mathbb{R})$ is of type A_{n-1}, and each positive root corresponds to a particular entry of the matrix. The simple roots correspond to the super-diagonal entries of the matrix. Hence we use the term "simple case" to describe the situation in which we insist that the super-diagonal, or simple, entries of u and v all avoid a neighbourhood of 0.

The process works especially well in $\mathrm{SL}_n(\mathbb{R})$. As mentioned in the introduction we consider only those semisimple real Lie groups whose Lie algebras are split. This is because we require that the root spaces $\mathfrak{g}_\lambda$ all have dimension 1.

The ideas laid out below for the simple case could potentially be extended to a more general situation where the simple entries are allowed to be zero. First one should find the largest parabolic subgroup of G which contains u in its unipotent radical. When looking at the root system, this corresponds to taking bites out of it — i.e. removing the linear spans of certain simple roots. We would then need to find a new subset of the set of positive roots which can play the role of the simple

roots. Then conjugate by an element of the parabolic subgroup in order to ensure the corresponding entries are non-zero.

Let N be a maximal unipotent subgroup of G. As a consequence of the root-space decomposition of the Lie algebra $\mathfrak{g}$, associated to G is a reduced root system Λ, containing a subset Λ^+ of positive roots, such that any element $u \in N = \exp(\mathfrak{n})$ can be uniquely expressed as

$$(3) \qquad u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right)$$

where each Y_λ lies in a root-space $\mathfrak{g}_\lambda$ in $\mathfrak{g}$. Based on this, we introduce some terminology. For u as above, the element Y_λ will be called the λ–*entry* of u. If λ is a simple root in Λ^+, that is $\lambda \in \Pi$, then we will say that Y_λ is a *simple entry*.

6.1. **Outline of the method.** The idea is to conjugate $u \in N$ by a sequence elements of the form $\exp(Z_\mu)$, where $Z_\mu \in \mathfrak{g}_\mu$, or by a commutator of two such elements (from two distinct root-spaces), each step in the sequence removing a λ–entry of u.

For example, when dealing with unipotent upper triangular matrices in $\mathrm{SL}_n(\mathbb{R})$ each entry in the triangle above the diagonal corresponds to a root. The simple roots correspond to the super-diagonal entries, that is those which lie adjacent to the diagonal. Take

$$u := \begin{pmatrix} 1 & x_1 & y_1 & z_1 \\ 0 & 1 & x_2 & y_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Here the x_i entries are the simple entries of u, the y_i terms correspond to roots of height 2 and z_1 to the unique root of height 3. Suppose all these entries are non-zero. We will conjugate u by an elementary matrix to make the y_1 term vanish:

$$\begin{pmatrix} 1 & \alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} u \begin{pmatrix} 1 & -\alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & x_1 & y_1 + \alpha x_2 & z_1 + \alpha y_2 \\ 0 & 1 & x_2 & y_2 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

So if we set $\alpha = -\frac{y_1}{x_2}$ then the entry where y_1 was has now been made to be zero. Notice that all simple entries and the other entry of height 2 are unchanged by this conjugation — the only collateral damage is to entries corresponding to roots of strictly greater height that the entry we removed.

The idea is to repeat this process, next removing the other height 2 entry. This will again cause collateral damage, but it will similarly only effect the height 3 entry. This then is the last entry to be removed and is done so by one last conjugation, but in this case there is no root of greater height than 3, so there will be no collateral damage.

We have conjugated u to

$$u' := \begin{pmatrix} 1 & x_1 & 0 & 0 \\ 0 & 1 & x_2 & 0 \\ 0 & 0 & 1 & x_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

via upper triangular matrices with rational entries. We do the same for another upper triangular matrix v, reducing to v' in a similar manner. If u and v are conjugate in $\mathrm{SL}_n(\mathbb{Z})$ then u' and v' must be conjugate in $\mathrm{SL}_n(\mathbb{Q})$. In fact, when

the simple entries in u and v are positive we can find a diagonal matrix over $\mathbb{R}$ to do the job:

$$
\begin{pmatrix}
\alpha_1 & 0 & 0 & 0 \\
0 & \alpha_2 & 0 & 0 \\
0 & 0 & \alpha_3 & 0 \\
0 & 0 & 0 & \alpha_4
\end{pmatrix}
u'
\begin{pmatrix}
\alpha_1 & 0 & 0 & 0 \\
0 & \alpha_2 & 0 & 0 \\
0 & 0 & \alpha_3 & 0 \\
0 & 0 & 0 & \alpha_4
\end{pmatrix}^{-1}
=
\begin{pmatrix}
1 & w_1 & 0 & 0 \\
0 & 1 & w_2 & 0 \\
0 & 0 & 1 & w_3 \\
0 & 0 & 0 & 1
\end{pmatrix}
$$

where $\alpha_1^4 = \frac{w_1^3 w_2^2 w_3}{x_1^3 x_2^2 x_3}$, $\alpha_2^4 = \frac{x_1 w_2^2 w_3}{w_1 x_2^2 x_3}$, $\alpha_3^4 = \frac{x_1 x_2^2 w_3}{w_1 w_2^2 x_3}$, and $\alpha^4 = \frac{x_1 x_2^2 x_3}{w_1 w_2^2 w_3}$.

From our point of view, the crucial aspect of this process is that we can keep track of the size of the conjugator in each step and also control the extent of the collateral damage occurring to entries of greater height.

6.2. **The metric on G.** The Killing form on $\mathfrak{g}$ is the symmetric bilinear form $B : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ given by $B(V, W) = \mathrm{Trace}(\mathrm{ad}(V)\mathrm{ad}(W))$. Given $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, the Cartan decomposition at $p \in X$, the Killing form is negative definite on $\mathfrak{k}$ and positive definite on $\mathfrak{p}$. Furthermore $\mathfrak{p}$ is the orthogonal complement of $\mathfrak{k}$ with respect to the Killing form. Let θ_p be the Cartan involution on $\mathfrak{g}$ defined at p, that is θ_p acts on $\mathfrak{k}$ as the identity and $\theta_p(Y) = -Y$ for any $Y \in \mathfrak{p}$. We can then define an inner product on $\mathfrak{g}$ as follows (see [Ebe96, §2.7] or [Hel01, Ch. III Prop 7.4]):

$$
\varphi_p(Y, Z) = -B(\theta_p Y, Z), \text{ for all } Y, Z \in \mathfrak{g}.
$$

This then determines a left-invariant Riemannian metric d_G on G. Denote the norm on $\mathfrak{g}$ corresponding to φ_p by $\|.\|$.

We will be interested in the effect of the Lie bracket on the size of elements from the root-spaces of $\mathfrak{g}$.

Proposition 6.1. *Suppose $\mathfrak{g}$ is a split real Lie algebra. Let $Y_\lambda \in \mathfrak{g}_\lambda$ and $Y_\mu \in \mathfrak{g}_\mu$, where $\lambda \in \Pi$ and $\mu \in \Lambda^+$ and $\lambda + \mu$ is a root. Then there exist constants $c_1 \geq c_0 > 0$, independent of the choice of $Y_\lambda, Y_\mu, \lambda, \mu$, such that*

$$
c_1 \|Y_\lambda\| \|Y_\mu\| \geq \|[Y_\mu, Y_\lambda]\| \geq c_0 \|Y_\lambda\| \|Y_\mu\|.
$$

Proof. This follows from the fact that the root-spaces have dimension one and also from the bilinearity of the Lie bracket and of the inner product φ_p. In particular, if we let Z_λ denote one of the two elements of $\mathfrak{g}_\lambda$ such that $\|Z_\lambda\| = 1$, and similarly for Z_μ, then for $\alpha, \beta \in \mathbb{R}$ such that $Y_\lambda = \alpha Z_\lambda$ and $Y_\mu = \beta Z_\mu$:

$$
\|[Y_\lambda, Y_\mu]\| = |\alpha| \, |\beta| \, c_{\lambda, \mu}
$$

where $c_{\lambda, \mu} = \|[Z_\lambda, Z_\mu]\|$. By taking

$$
c_0 = \min\{c_{\lambda, \mu} \mid \mu \in \Lambda^+, \lambda \in \Pi \text{ such that } \lambda + \mu \in \Lambda^+\}
$$
$$
c_1 = \max\{c_{\lambda, \mu} \mid \mu \in \Lambda^+, \lambda \in \Pi \text{ such that } \lambda + \mu \in \Lambda^+\}
$$

we obtain the result, since $|\alpha| = \|Y_\lambda\|$ and $|\beta| = \|Y_\mu\|$. □

The following tells us that the size of any λ–entry of u will give us a lower bound for the size of $d_G(1, u)$.

Lemma 6.2. *Let $u \in N$ be as in (3). For each $\lambda \in \Lambda^+$ we have $\|Y_\lambda\| \leq d_G(1, u)$.*

Proof. Since the root-spaces $\mathfrak{g}_\lambda$ for $\lambda \in \Lambda^+$ are pairwise orthogonal with respect to the Killing form, and hence also the inner product φ_p, we observe that:

$$
\|Y_\lambda\| \leq \left\| \sum_{\lambda \in \Lambda} Y_\lambda \right\| = d_G(1, u).
$$

□

7. The role of the root system

7.1. **Relating the root system to conjugation.** If N is a maximal unipotent subgroup with corresponding root system Λ then any element in N can be written uniquely as

$$(4) \qquad u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda \right)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$. We begin by studying the behaviour of the λ–entries under the action of conjugation by elements in N of the form $\exp(Z_\mu)$, where $Z_\mu \in \mathfrak{g}_\mu$, or a commutator of two such elements.

Lemma 7.1. *Let $u \in N$ be as in (4) and let $Z_\mu \in \mathfrak{g}_\mu$ for some $\mu \in \Lambda^+$. When we conjugate u by $\exp(Z_\mu)$ all entries of u are unchanged except (possibly) for the λ–entries where $\lambda = r\mu + \lambda'$ for some $r \in \mathbb{N}$ $(r \neq 0)$ and $\lambda' \in \Lambda^+$ such that $Y_{\lambda'} \neq 0$. Furthermore, the λ–entry of $\exp(Z_\mu)u\exp(-Z_\mu)$ is*

$$\sum_{r\mu+\lambda'=\lambda} \frac{(\mathrm{ad}Z_\mu)^r Y_{\lambda'}}{r!}$$

where the sum takes values of r from $\mathbb{N} \cup \{0\}$ and λ' from Λ^+.

Proof. Let $Z = Z_\mu$. We observe that:

$$\begin{aligned}
\exp(Z)u\exp(-Z) &= \exp\left(e^{\mathrm{ad}(Z)} \sum_{\lambda' \in \Lambda^+} Y_{\lambda'} \right) \\
&= \exp\left(\sum_{r=0}^{\infty} \sum_{\lambda' \in \Lambda^+} \frac{(\mathrm{ad}Z)^r Y_{\lambda'}}{r!} \right)
\end{aligned}$$

Recall that if $r\mu + \lambda'$ is not a root then $(\mathrm{ad}Z)^r Y_{\lambda'} = 0$. Otherwise $(\mathrm{ad}Z)^r Y_{\lambda'} \in \mathfrak{g}_{r\mu+\lambda'}$. It follows that if λ cannot be written as $r\mu + \lambda'$ for any $r \neq 0$ or any λ' then the λ–entry of u in unchanged by this conjugation. $\qquad\square$

The preceding Proposition is important in recognising the link between conjugation of unipotent elements and the root system of G. In particular we can see that if the λ–entry of u is affected by conjugating by $\exp(Z_\mu)$ then the height of λ, denoted $\mathrm{ht}\lambda$, must be greater than $\mathrm{ht}\mu$. Furthermore, the affected entries whose height is precisely $\mathrm{ht}\mu + 1$ will be in the set $\{\mu\} + \Pi$, where Π is the set of simple roots in Λ^+. This is crucial for motivating Lemma 7.4.

To complete the picture which lies behind the scenes of Lemma 7.4 we must also consider conjugating by a commutator of two elements. Building up to this, which is Lemma 7.3, we give the following:

Lemma 7.2. *Let $u \in N$ be as in (4) and let $Z_1 \in \mathfrak{g}_{\mu_1}$, $Z_2 \in \mathfrak{g}_{\mu_2}$ for some $\mu_1, \mu_2 \in \Lambda^+$. When we conjugate u by $\exp(Z_1)\exp(Z_2)$ all entries of u are unchanged except (possibly) for the λ–entries where $\lambda = r\mu_1 + t\mu_2 + \lambda'$ for some $\lambda' \in \Lambda^+$ and non-negative integers r, t where at least one of r, t is non-zero.*

Proof. As in the proof of Lemma 7.1 we get:

$$\exp(Z_2)\exp(Z_1)u\exp(-Z_1)\exp(-Z_2) = \sum_{t=0}^{\infty}\sum_{r=0}^{\infty}\sum_{\lambda' \in \Lambda^+} \frac{(\mathrm{ad}Z_2)^t(\mathrm{ad}Z_1)^r}{r!t!} Y_{\lambda'}$$

Since $(\mathrm{ad}Z_2)^t(\mathrm{ad}Z_1)^r Y_{\lambda'} \in \mathfrak{g}_{r\mu_1+t\mu_2+\lambda'}$ if $r\mu_1+t\mu_2+\lambda'$ is a root, or is zero otherwise, it follows as in Lemma 7.1 that if λ cannot be expressed as $r\mu_1 + t\mu_2 + \lambda'$ for some λ' and non-negative integers r, t where one of r, t is non-zero, then the λ–entry of u in not affected by this conjugation process. $\qquad\square$

Rather than conjugating by $\exp(Z_1)\exp(Z_2)$ we will conjugate by their commutator $[\exp(Z_1)\exp(Z_2)]$. The extra terms in the product act to clean up any effect conjugating by $\exp(Z_1)\exp(Z_2)$ had on the entries of height less than or equal to $\operatorname{ht}\mu_1 + \operatorname{ht}\mu_2$. Observe that the λ–entry of the conjugate of u by $\exp(Z_1)\exp(Z_2)$ is:

$$\sum_{r\mu_1+t\mu_2+\lambda'=\lambda} \frac{(\operatorname{ad}Z_2)^t(\operatorname{ad}Z_1)^r}{r!t!} Y_{\lambda'}.$$

Suppose $\operatorname{ht}\lambda \leq \operatorname{ht}\mu_1 + \operatorname{ht}\mu_2$. Then in each term in the sum either $r = 0$ or $t = 0$. We can therefore rewrite it as:

$$\sum_{r\mu_1+\lambda'=\lambda} \frac{(\operatorname{ad}Z_1)^r}{r!} Y_{\lambda'} + \sum_{t\mu_2+\lambda'=\lambda} \frac{(\operatorname{ad}Z_2)^t}{t!} Y_{\lambda'} - Y_{\lambda}.$$

The extra $-Y_\lambda$ term is needed because when $r = t = 0$ we count Y_λ twice when it should only be counted once. Next we conjugate by $\exp(-Z_1)\exp(-Z_2)$ and we get the following for the λ–entry:

$$\sum_{R\mu_1+\lambda'=\lambda} \frac{(\operatorname{ad}(-Z_1))^R}{R!} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(\operatorname{ad}Z_1)^r}{r!} Y_{\lambda''} + \sum_{t\mu_2+\lambda''=\lambda'} \frac{(\operatorname{ad}Z_2)^t}{t!} Y_{\lambda''} - Y_{\lambda'} \right)$$

$$+ \sum_{T\mu_2+\lambda'=\lambda} \frac{(\operatorname{ad}(-Z_2))^T}{T!} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(\operatorname{ad}Z_1)^r}{r!} Y_{\lambda''} + \sum_{t\mu_2+\lambda''=\lambda'} \frac{(\operatorname{ad}Z_2)^t}{t!} Y_{\lambda''} - Y_{\lambda'} \right) - Y_{\lambda}$$

Since $\operatorname{ht}\lambda \leq \operatorname{ht}\mu_1 + \operatorname{ht}\mu_2$, we cannot write $\lambda = R\mu_1 + t\mu_2 + \lambda'$ when both R, t are non-zero (and similarly for r and T). Hence this expression can be reduced to:

$$\sum_{R\mu_1+\lambda'=\lambda} \frac{(\operatorname{ad}(-Z_1))^R}{R!} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(\operatorname{ad}Z_1)^r}{r!} Y_{\lambda''} \right)$$

$$+ \sum_{T\mu_2+\lambda'=\lambda} \frac{(\operatorname{ad}(-Z_2))^T}{T!} \left(\sum_{t\mu_2+\lambda''=\lambda'} \frac{(\operatorname{ad}Z_2)^t}{t!} Y_{\lambda''} \right) - Y_{\lambda}$$

This can be rewritten as:

$$\sum_{R\mu_1+\lambda'=\lambda} \left(\sum_{r\mu_1+\lambda''=\lambda'} \frac{(-1)^R}{R!r!} (\operatorname{ad}Z_1)^{R+r} Y_{\lambda''} \right)$$

$$+ \sum_{T\mu_2+\lambda'=\lambda} \left(\sum_{t\mu_2+\lambda''=\lambda'} \frac{(-1)^T}{T!t!} (\operatorname{ad}Z_2)^{T+t} t! Y_{\lambda''} \right) - Y_{\lambda}$$

Notice that whenever $R + r \neq 0$ all the terms cancel, since if $R + r = k \neq 0$ then the coefficient of $(\operatorname{ad}Z_1)^k Y_{\lambda''}$ is:

$$\sum_{R+r=k} \frac{(-1)^R}{R!r!} = 0.$$

A similar statement holds for $T + t \neq 0$. Hence, whenever $\operatorname{ht}\lambda \leq \operatorname{ht}\mu_1 + \operatorname{ht}\mu_2$, the λ–entry is Y_λ. We use this in the following:

Lemma 7.3. *Let $u \in N$ be as in (4) and let $Z_1 \in \mathfrak{g}_{\mu_1}$, $Z_2 \in \mathfrak{g}_{\mu_2}$ for some $\mu_1, \mu_2 \in \Lambda^+$. When we conjugate u by $[\exp(Z_1), \exp(Z_2)]$ all entries of u are unchanged except (possibly) for the λ–entries where $\lambda = r\mu_1 + t\mu_2 + \lambda'$ for some $\lambda' \in \Lambda^+$ and non-negative integers r, t.*

Furthermore, for such λ, the λ–entry of the conjugate is

$$\sum_{(r+t)\mu_1+(s+u)\mu_2+\lambda'=\lambda} \frac{\mathrm{ad}(-Z_2)^u\mathrm{ad}(-Z_1)^t\mathrm{ad}(Z_2)^s\mathrm{ad}(Z_1)^rY'_\lambda}{r!s!t!u!}$$

where the summation takes non-negative integers r, s, t, u and positive roots λ'.

Proof. By repeating Lemma 7.1 we get that the λ–entry of the conjugate of u by $[\exp(Z_1), \exp(Z_2)]$ is given by

$$\sum_{(r+t)\mu_1+(s+u)\mu_2+\lambda'=\lambda} \frac{\mathrm{ad}(-Z_2)^u\mathrm{ad}(-Z_1)^t\mathrm{ad}(Z_2)^s\mathrm{ad}(Z_1)^rY'_\lambda}{r!s!t!u!}$$

as required.

This, together with the argument preceding the statement of the Proposition, gives the result. $\qquad\square$

The important difference between Lemma 7.3 and Lemma 7.2 is that when we conjugate by the commutator all entries of height no more than $\mathrm{ht}\mu_1 + \mathrm{ht}\mu_2$ are left unchanged. Furthermore the only (possibly) affected entries of height $\mathrm{ht}\mu_1+\mathrm{ht}\mu_2+1$ are precisely those entries corresponding to roots in the set $\{\mu_1\} + \{\mu_2\} + \Pi$ where Π is the set of simple roots in Λ^+.

7.2. **An ordering on the root system.** We consider in this paper the "simple case," by which we mean the case when u is given by

$$u = \exp\left(\sum_{\lambda\in\Lambda^+} Y_\lambda\right)$$

and $Y_\lambda \neq 0$ for each simple root λ. The aim is to find a sequence of elements like those considered in Lemmas 7.1 and 7.3 which reduce u to a form where the only non-zero λ–entries are those where λ is simple. The following Lemma is necessary to ensure that such a sequence of elements can be found in the simple case.

Lemma 7.4. *Let Λ^+ be a set of positive roots and Π the corresponding simple roots associated to a reduced root system Λ. We can assign to Λ^+ an ordering, which we will denote by $<$, such that for every $\lambda \in \Lambda^+ \setminus \Pi$ either:*

(a) *there exists some root μ such the set $\{\mu\}+\Pi$ contains λ and $\lambda \neq \lambda' \in \{\mu\}+\Pi$ implies $\lambda < \lambda'$; or*

(b) *there exist roots $\mu_1, \mu_2 \in \Lambda^+$ such that $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda\}$ and $\mu_1 + \mu_2$ is not a root.*

Remark: Case (a) corresponds to conjugation by something in $\exp(\mathfrak{g}_\mu)$, see Lemma 7.1. Case (b) corresponds to conjugating by a commutator as in Lemma 7.3. This Lemma, combined with Lemmas 7.1 and 7.3, tells us that we can always conjugate $u \in N$ by an element of N in such a way that we can choose the smallest entry of u which is affected by the conjugation.

Proof of Lemma 7.4. Before we proceed, note that if we find μ_1, μ_2 satisfying (b) but $\mu_1 + \mu_2$ is a root, then case (a) also applies.

Suppose that Λ is the sum of irreducible root systems $\Lambda_1, \ldots, \Lambda_r$ and that $\Lambda^+ = \Lambda_1^+ \cup \ldots \cup \Lambda_r^+$. Suppose also that on each Λ_i^+ we have an ordering $<_i$ which satisfies the Lemma. Then we can define an ordering $<$ on Λ^+ given by $\lambda < \mu$ if and only if

(1) $\lambda \in \Lambda_i^+$ and $\mu \in \Lambda_j^+$ such that $i < j$; or

(2) if $\lambda, \mu \in \Lambda_i^+$ for some i then $\lambda <_i \mu$.

If λ and μ are in different irreducible root systems inside Λ, then $\lambda + \mu$ cannot be a root. Hence it follows that if $<_i$ satisfies the Lemma for each i, then so does $<$. Thus it suffices to check the conditions of the Lemma for each irreducible root system.

In the classical root systems A_n, B_n, C_n, D_n, we make the base assumption that the simple roots are $\Pi = \{\lambda_1, \ldots, \lambda_n\}$ and are ordered by $\lambda_1 > \lambda_2 > \ldots > \lambda_n$. Note that because we will assume that the Lie algebra is split we do not need to consider BC_n root systems.

ROOT SYSTEMS OF TYPE A_n:

This is the root system associated to $\mathrm{SL}_n(\mathbb{Z})$ so we expect this to be straightforward. The non-simple positive roots will be sums of consecutive simple roots:

$$\lambda_i + \lambda_{i+1} + \ldots + \lambda_j$$

for $1 \le i < j \le n$. The ordering we assign is in two steps: primarily we order by height, then within each height we order the elements lexicographically. So if $\lambda = \lambda_i + \ldots + \lambda_j$ then we take $\mu = \lambda_i + \ldots + \lambda_{j-1}$. It follows that:

$$\{\mu\} + \Pi = \left\{ \begin{array}{ll} \{\lambda, \lambda_{i-1} + \ldots + \lambda_{j-1}\} & \text{if } i \ne 1 \\ \{\lambda\} & \text{if } i = 1 \end{array} \right.$$

and hence our chosen ordering satisfies the requirements of the lemma.

ROOT SYSTEMS OF TYPE B_n:

The non-simple positive roots are of the following forms:

$$\begin{array}{ll} \lambda_i + \ldots + \lambda_j & \text{for } 1 \le i < j \le n \\ \lambda_i + \ldots + \lambda_{j-1} + 2\lambda_j + \ldots + 2\lambda_n & \text{for } 1 \le i < j \le n \end{array}$$

We order the roots as we did for type A_n: first by height, then order the elements of each height lexicographically. If we first take λ of the first form listed above, i.e. $\lambda = \lambda_i + \ldots + \lambda_j$. Then we take $\mu = \lambda_i + \ldots + \lambda_{j-1}$ and observe that:

$$\{\mu\} + \Pi = \left\{ \begin{array}{ll} \{\lambda, \lambda_{i-1} + \ldots + \lambda_{j-1}\} & \text{if } i \ne 1 \\ \{\lambda\} & \text{if } i = 1 \end{array} \right.$$

satisfies the required conditions. If on the other hand we consider

$$\lambda = \lambda_i + \ldots + \lambda_{j-1} + 2\lambda_j + \ldots + 2\lambda_n$$

then we take

$$\mu = \left\{ \begin{array}{ll} \lambda_i + \ldots + \lambda_j + 2\lambda_{j+1} + \ldots + 2\lambda_n & \text{if } j \ne n \\ \lambda_i + \ldots + \lambda_n & \text{if } j = n \end{array} \right.$$

and observe that:

$$\{\mu\} + \Pi = \left\{ \begin{array}{ll} \{\lambda, \lambda_{i-1} + \ldots + \lambda_j + 2\lambda_{j+1} + \ldots + 2\lambda_n\} & \text{if } i \ne 1, j \ne n \\ \{\lambda, \lambda_{i-1} + \ldots + \lambda_n\} & \text{if } i \ne 1, j = n \\ \{\lambda\} & \text{if } i = 1 \end{array} \right.$$

satisfies the requirements for every choice of i, j.

ROOT SYSTEMS OF TYPE C_n:

The positive non-simple roots have one of the following forms:

$$\begin{array}{ll} \lambda_i + \ldots + \lambda_j & \text{for } 1 \le i < j \le n \\ 2\lambda_i + \ldots + 2\lambda_{n-1} + \lambda_n & \text{for } 1 \le i \le n-1 \\ \lambda_i + \ldots + \lambda_{j-1} + 2\lambda_j + \ldots + 2\lambda_{n-1} + \lambda_n & \text{for } 1 \le i < j \le n-1 \end{array}$$

We order these first by height, then order the elements of the same height by lexicographic ordering. We now give the choice for μ in each case.

First, for $1 \le i < j \le n$, let $\lambda = \lambda_i + \ldots + \lambda_j$. Then we take $\mu = \lambda_i + \ldots + \lambda_{j-1}$ and we have:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \ldots + \lambda_{j-1}\} & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

Under our chosen ordering these satisfy the requirements of the Lemma.

Second, for $1 \le i \le n-1$, let $\lambda = 2\lambda_i + \ldots + 2\lambda_{n-1} + \lambda_n$. Then we take $\mu = \lambda_i + 2\lambda_{i+1} + \ldots + 2\lambda_{n-1} + \lambda_n$ if $i \neq n-1$ or $\mu = \lambda_{n-1} + \lambda_n$ if $i = n-1$ and we have:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \lambda_i + 2\lambda_{i+1} \ldots + 2\lambda_{n-1} + \lambda_n\} & \text{if } i \neq n-1 \text{ and } i \neq 1 \\ \{\lambda, \lambda_{n-2} + \lambda_{n-1} + \lambda_n\} & \text{if } i = n-1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

In each case the elements of $\{\mu\} + \Pi$ are at least as big as λ in our chosen ordering, so the Lemma is satisfied in this case.

Finally, for $1 \le i < j \le n-1$, let $\lambda = \lambda_i + \ldots + \lambda_{j-1} + 2\lambda_j + \ldots + 2\lambda_{n-1} + \lambda_n$. We take $\mu = \lambda_i + \ldots + \lambda_j + 2\lambda_{j+1} + \ldots + 2\lambda_{n-1} + \lambda_n$ if $j \neq n-1$ or $\mu = \lambda_i + \ldots + \lambda_n$ when $j = n-1$. Then:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \ldots + \lambda_j + 2\lambda_{j+1} \ldots + 2\lambda_{n-1} + \lambda_n\} & \text{if } j \neq n-1 \text{ and } i \neq 1 \\ \{\lambda, \lambda_{i-1} + \ldots + \lambda_n\} & \text{if } j = n-1 \text{ and } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

The requirements of the Lemma are satisfied in each case, and hence it follows that the Lemma holds for root systems of type C_n.

ROOT SYSTEMS OF TYPE D_n:

The non-simple positive roots in the root system D_n are of one of the following two types:

$$\begin{array}{ll} \lambda_i + \ldots + \lambda_{j-1} & \text{if } 1 \le i < j \le n \\ \lambda_i + \ldots + \lambda_{n-2} + \lambda_j + \ldots + \lambda_n & \text{if } 1 \le i < j \le n \end{array}$$

Apply the same ordering to D_n as we applied to each of the preceding root systems: first order by height, then order the elements of the same height lexicographically. In most instances we are able to satisfy the conditions of the Lemma by choosing a single $\mu \in \Lambda^+$. However there are some for which we must use the second allowable case, namely find two positive roots μ_1, μ_2 to satisfy the Lemma.

We first suppose $\lambda = \lambda_i + \ldots + \lambda_{j-1}$ where $1 \le i < j < n$. Then we take $\mu = \lambda_i + \ldots \lambda_{j-2}$ and observe:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \ldots + \lambda_{j-2}\} & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

Hence the conditions of the Lemma are satisfied in each case.

For $1 \le i < j \le n$ let $\lambda = \lambda_i + \ldots + \lambda_{n-2} + \lambda_j + \ldots + \lambda_n$. First assume $i \neq n-2$ and $j \neq n-1, n$. If we take $\mu = \lambda_i + \ldots + \lambda_{n-2} + \lambda_{j+1} + \ldots + \lambda_n$ then:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \ldots + \lambda_{n-2} + \lambda_{j+1} + \ldots + \lambda_n & \text{if } i \neq 1 \\ \{\lambda\} & \text{if } i = 1 \end{cases}$$

and the Lemma is satisfied.

Now suppose $j = n$, then $\lambda = \lambda_i + \ldots + \lambda_{n-2} + \lambda_n$. Take $\mu = \lambda_i + \ldots + \lambda_{n-2}$ then:

$$\{\mu\} + \Pi = \begin{cases} \{\lambda, \lambda_{i-1} + \ldots + \lambda_{n-2}, \lambda_i + \ldots + \lambda_{n-1}\} & \text{if } i \neq 1 \\ \{\lambda, \lambda_1 + \ldots + \lambda_{n-1}\} & \text{if } i = 1 \end{cases}$$

and our choice of μ here satisfies the requirements of the Lemma.

We are left with the cases when $\lambda = \lambda_i + \ldots + \lambda_{n-1}$ and when $\lambda = \lambda_{n-2} + \lambda_{n-1} + \lambda_n$. In the former case we take $\mu_1 = \lambda_i + \ldots + \lambda_{n-3}$ and $\mu_2 = \lambda_{n-1}$ and observe

the only way to make a root by adding μ_1, μ_2 and a simple root together is if the simple root is λ_{n-2}, thus giving λ. Hence $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda\}$. In the latter case we take $\mu_1 = \lambda_{n-1}$ and $\mu_2 = \lambda_n$. Similarly, since the only simple root which we can add to $\mu_1 + \mu_2$ and still have a root is λ_{n-2}, we have $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda\}$.

This completes the verification of the Lemma in the case when the root system is of type D_n.

ROOT SYSTEMS OF TYPE E_6, E_7, E_8, F_4, G_2:

These are dealt with in the appendix. For root systems E_8 and F_4 a table is produced with an example of an ordering satisfying the Lemma. They also give suitable choices of μ or of μ_1 and μ_2 for each non-simple positive root. Table 1 gives the ordering for E_8, and hence for E_7 and E_6 by using the induced ordering. Table 2 gives the ordering for F_4. Figures 9, 10 and 11 provide a visual method of checking in each case that the given root μ satisfies the requirements: given $\mu \in \Lambda^+$ one can quickly see what $\{\mu\} + \Pi$ will be by following all edges heading down the page from μ to the row below.

When dealing with G_2, there is only one root of each height strictly greater than 1, hence we can order the roots by height alone.   □

8. CONSTRUCTION OF A SHORT CONJUGATOR

8.1. **Reduction of the simple case.** From here on in we will assume that $\mathfrak{g}$ is a split real Lie algebra, meaning that the root spaces $\mathfrak{g}_\lambda$ are 1–dimensional. We first give an algorithm to reduce $u \in N$, all of whose simple entries are non-zero, to $u' \in N$, all of whose non-simple entries are zero and the simple entries of u' are equal to those of u. Write u in terms of the elements from the root-spaces of $\mathfrak{g}$:

$$u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right)$$

where $Y_\lambda \in \mathfrak{g}_\lambda$ for each $\lambda \in \Lambda^+$. Assign to Λ^+ the ordering from Lemma 7.4. The algorithm is based on an iteration of the following result:

Lemma 8.1. *Let λ_0 be the smallest non-simple root such that Y_{λ_0} is non-zero. Then there exists $g \in N$ and a positive constant $c_0 > 0$ such that:*

(i) *the λ_0–entry of gug^{-1} is zero and all entries corresponding to smaller roots are unchanged; and*

(ii) *$d_G(1, g) \leq \frac{\|Y_{\lambda_0}\|}{c_0 \delta}$, where $\delta = \min\{\|Y_{\lambda_i}\| \mid \lambda_i \in \Pi\}$.*

Proof. We begin by applying Lemma 7.4 to λ_0. This gives us either:

(a) $\mu \in \Lambda^+$ such that λ_0 is minimal in $\{\mu\} + \Pi$; or

(b) $\mu_1, \mu_2 \in \Lambda^+$ such that $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda_0\}$ and $\mu_1 + \mu_2$ is not a root.

First suppose (a) holds. Take $g = \exp(Z_\mu)$ where $Z_\mu \in \mathfrak{g}_\mu$ is chosen so that

$$[Z_\mu, Y_{\lambda_i}] = -Y_{\lambda_0}$$

where λ_i is the simple root such that $\mu + \lambda_i = \lambda_0$. By Lemma 7.1, the λ_0–entry of gug^{-1} is, by construction,

$$Y_{\lambda_0} + \mathrm{ad}(Z_\mu)Y_{\lambda_i} = 0$$

and the other affected entries are of the form $r\mu + \lambda$ for some $\lambda \in \Lambda^+$. All of these are larger than λ_0 in the ordering from Lemma 7.4, hence the first part of the lemma is proved when case (a) holds.

Now suppose that instead case (b) holds. Then we take $g = [\exp(Z_1), \exp(Z_2)]$ where $Z_i \in \mathfrak{g}_{\mu_i}$ for $i = 1, 2$. By Lemma 7.3, the λ_0–entry of gug^{-1} is

$$\sum_{(r+t)\mu_1 + (s+u)\mu_2 + \lambda' = \lambda_0} \frac{\mathrm{ad}(-Z_2)^u \mathrm{ad}(-Z_1)^t \mathrm{ad}(Z_2)^s \mathrm{ad}(Z_1)^r Y'_\lambda}{r! s! t! u!}$$

where the summation takes non-negative integers r, s, t, u and positive roots λ'. Since $Y_\lambda = 0$ for non-simple roots $\lambda < \lambda_0$, there is no other way to obtain a non-zero term in the sum except by either taking $r = s = t = u = 0$ and $\lambda' = \lambda_0$ or with $\lambda' = \lambda_i \in \Pi$ such that $\mu_1 + \mu_2 + \lambda_i = \lambda_0$. In the latter case we know $r + t = 1 = s + u$. Hence there are only finitely many combinations to consider and the λ_0 entry becomes:

$$\mathrm{ad}(Z_2)\mathrm{ad}(Z_1)Y_{\lambda_i} + \mathrm{ad}(-Z_2)\mathrm{ad}(Z_1)Y_{\lambda_i} +$$
$$\mathrm{ad}(-Z_1)\mathrm{ad}(Z_2)Y_{\lambda_i} + \mathrm{ad}(-Z_2)\mathrm{ad}(-Z_1)Y_{\lambda_i} + Y_{\lambda_0}$$

which simplifies to

$$\mathrm{ad}(Z_2)\mathrm{ad}(Z_1)Y_{\lambda_i} - \mathrm{ad}(Z_1)\mathrm{ad}(Z_2)Y_{\lambda_i} + Y_{\lambda_0}.$$

Finally, by application of the Jacobi identity, we see this is equal to

$$[[Z_2, Z_1], Y_{\lambda_i}] + Y_{\lambda_0}.$$

Hence, by choosing Z_1 and Z_2 so that $[[Z_2, Z_1], Y_{\lambda_i}] = -Y_{\lambda_0}$, the λ_0–entry of gug^{-1} is zero.

Finally, Lemma 7.3 tells us that entries corresponding to roots of height less than or equal to $\mathrm{ht}\mu_1 + \mathrm{ht}\mu_2$ are unchanged. Since also $\{\mu_1\} + \{\mu_2\} + \Pi = \{\lambda_0\}$, all entries corresponding to roots smaller than λ_0 are unaffected. Thus we have proved (i).

Note that we have the flexibility to choose Z_μ, Z_1 and Z_2 as above because each root-space has dimension one so we only need to choose the appropriate scalar multiple of a basis element to get what we want.

Now we look at the size of g. If $g = \exp(Z_\mu)$ arises from a situation like (a) then, since we chose Z_μ to satisfy $[Z_\mu, Y_{\lambda_i}] = -Y_{\lambda_0}$, we can use Proposition 6.1 to show:

$$d_G(1, g) = \|Z_\mu\| \leq \frac{\|Y_{\lambda_0}\|}{c_0 \|Y_{\lambda_i}\|} \leq \frac{\|Y_{\lambda_0}\|}{c_0 \delta}.$$

Suppose instead that $g = [\exp(Z_1), \exp(Z_2)]$, as is necessary for case (b). Using the Baker–Campbell–Hausdorff formula, $g = \exp([Z_1, Z_2])$ since $\mu_1 + \mu_2$ is a not a root. Then, again using Proposition 6.1 and our choice of Z_1, Z_2 such that $[[Z_2, Z_1], Y_{\lambda_i}] = -Y_{\lambda_0}$, we see that:

$$d_G(1, g) = \|[Z_1, Z_2]\| \leq \frac{\|Y_{\lambda_0}\|}{c_0 \|Y_{\lambda_i}\|} \leq \frac{\|Y_{\lambda_0}\|}{c_0 \delta}.$$

This completes (ii). □

The following algorithm describes a process by which, in the simple case, we can reduce $u \in N$ to $u' \in N$, where u' has no non-simple entries.

Algorithm A. Let $u \in N$ be given by

$$u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right), \quad Y_\lambda \neq 0 \text{ for } \lambda \in \Pi.$$

We define a sequence of elements $u(i) \in N$ where $u(0) = u$ and $u(i+1)$ has one fewer non-zero non-simple entry than $u(i)$ and is obtained by

$$u(i) := g(i)u(i-1)g(i)^{-1}, \text{ for } i \geq 1$$

where $g(i)$ is determined by Lemma 8.1. This process clearly terminates as Λ^+ is a finite set. Let $g(1), \ldots, g(r)$ be the complete set of conjugators obtained. Define $g := g(r) \ldots g(1)$. Then $u' := u(r) = gug^{-1}$, which has no non-zero non-simple entries.

8.2. **The collateral damage of Algorithm A.** Suppose that the simple entries of u are bounded away from zero. In particular, define a function $\Delta : N \to [0, \infty)$ by

$$\Delta(u) = \min\{\|Y_{\lambda_i}\| \mid \lambda_i \in \Pi\}$$

and suppose there exists some $\delta > 0$ such that $\Delta(u) \geq \delta$. Note that Δ can be extended to all unipotent elements of G. This function measures, in some vague sense, the distance of u from the simple root-spaces of $\mathfrak{g}$.

Before determining the size of a short conjugator in G we need to determine the effect each step of Algorithm A has on the entries of u. This is a notion we described in Section 6.1 as collateral damage. We showed in Lemmas 7.1 and 7.3 that while removing the λ_0 entry of u it was possible that some of the entries of greater height could be altered in the process. We will call those entries affected by one of the steps of Algorithm A, other than the intended target entry, the *collateral damage* of this step.

In general we expect collateral damage. We can, nonetheless, use an iterative method, bounding the size of each $u(i)$ in the sequence. By applying Lemmas 8.1 and 6.2 we see that the first conjugator $g(1)$ will satisfy

$$(5) \qquad d_G(1, g(1)) \leq \frac{1}{c_0\delta} d_G(1, u).$$

The collateral damage of conjugating u by $g(1)$ includes elements of height greater than that of the smallest non-simple non-zero entry of u. Suppose $g(1), \ldots, g(t_1)$ correspond to the steps to remove all entries of height 2. Since conjugating by any of these will not effect any height 2 entry of u, each $g(i)$, for $1 \leq i \leq t_1$, will satisfy inequality (5) in place of $g(1)$. Let $g_{\mathrm{ht}}(2) = g(t_1) \ldots g(1)$. Then

$$d_G(1, g_{\mathrm{ht}}(2)) \leq \frac{R_2}{c_0\delta} d_G(1, u)$$

where R_2 is equal to the number of roots of height 2. After the first t_1 steps of Algorithm A we obtain an element $u_{\mathrm{ht}}(2) = g_{\mathrm{ht}}(2)ug_{\mathrm{ht}}(2)^{-1}$ whose entries of height 2 are all zero. Furthermore, by the triangle inequality

$$d_G(1, u_{\mathrm{ht}}(2)) \leq \left(\frac{2R_2}{c_0\delta} + 1\right) d_G(1, u).$$

Suppose the λ–entry of $u_{\mathrm{ht}}(2)$ is $Y_\lambda^{(2)}$. Then by Lemma 6.2

$$\|Y_\lambda^{(2)}\| \leq \left(\frac{2R_2}{c_0\delta} + 1\right) d_G(1, u).$$

By Lemma 8.1, the size of the next conjugator will be bounded above:

$$d_G(1, g(t_1 + 1)) \leq \frac{1}{c_0\delta}\left(\frac{2R_2}{c_0\delta} + 1\right) d_G(1, u)$$

noting that we can still use δ as defined above since the simple entries of $u_{\mathrm{ht}}(2)$ are exactly those of u. Let $g_{\mathrm{ht}}(3) = g(t_2) \ldots g(t_1 + 1)$, where $g(t_1 + 1), \ldots, g(t_2)$ are those conjugators from Algorithm A corresponding to the removal of height 3 entries of u. Then, as in the height 2 case, we get

$$d_G(1, g_{\mathrm{ht}}(3)) \leq \frac{R_3}{c_0\delta}\left(\frac{2R_2}{c_0\delta} + 1\right) d_G(1, u)$$

where R_3 is the number of roots of height 2. Then $u_{\mathrm{ht}}(3) = g_{\mathrm{ht}}(3)u_{\mathrm{ht}}(2)g_{\mathrm{ht}}(3)^{-1}$ has no entries of height 2 or 3, and it satisfies

$$d_G(1, u_{\mathrm{ht}}(3)) \leq \left(\frac{2R_3}{c_0\delta} + 1\right)\left(\frac{2R_2}{c_0\delta} + 1\right)d_G(1, u).$$

Continuing in this way, if r is the greatest height of a root in Λ^+, then for each $2 \leq i \leq r$ we have

$$d_G(1, g_{\mathrm{ht}}(i)) \leq \frac{R_i}{c_0\delta}\prod_{j=2}^{i-1}\left(\frac{2R_j}{c_0\delta} + 1\right)d_G(1, u).$$

Let $g = g_{\mathrm{ht}}(r)\ldots g_{\mathrm{ht}}(2)$. Then g is the element obtained from Algorithm A and conjugates u to an element u' whose non-simple entries are all zero, while its simple entries are the same as for u. Finally, we see that the size of g is bounded linearly by the size of u:

Proposition 8.2. *Let g be the conjugator obtained by Algorithm A such that the non-simple entries of gug^{-1} are all zero. Then*

$$d_G(1, g) \leq K(\delta)d_G(1, u)$$

where

$$K(\delta) = \sum_{i=2}^{r}\frac{R_i}{c_0\delta}\prod_{j=2}^{i-1}\left(\frac{2R_j}{c_0\delta} + 1\right).$$

8.3. **The last step towards finding a short conjugator.** Take u as above and let v be an element in N conjugate to u. Suppose we can express v as

$$\exp\left(\sum_{\lambda \in \Lambda^+} W_\lambda\right), \quad W_\lambda \in \mathfrak{g}_\lambda.$$

By applying Algorithm A we may assume that $Y_\lambda = 0 = W_\lambda$ for all non-simple roots $\lambda \in \Lambda^+$. Then, by choosing $H \in \mathfrak{a}$ appropriately, we can conjugate u to v using $\exp(H)$. To be precise:

$$
\begin{aligned}
gug^{-1} &= \exp(H)\exp\left(\sum_{\lambda \in \Pi} Y_\lambda\right)\exp(-H) \\
&= \exp\left(\sum_{\lambda \in \Pi} e^{\lambda(H)}Y_\lambda\right).
\end{aligned}
$$

Hence our choice of H needs to be such that $e^{\lambda(H)}Y_\lambda = W_\lambda$. We might ask, what if we need negative scalars? The following Proposition answers this question:

Proposition 8.3. *Let u and v be unipotent elements contained in the same maximal unipotent subgroup N of G. Suppose that u is conjugate to v in G and furthermore suppose that the non-simple entries of u and v are all trivial while the simple entries are all non-zero. Then there exists $H_0 \in \mathfrak{a}$ such that*

$$\exp(H_0)u\exp(-H_0) = v.$$

Proof. Let $g \in G$ be such that $gug^{-1} = v$. First observe that, since we are dealing with the simple case, both u and v fix the same unique chamber $\partial_\infty\mathcal{C}$ in the ideal boundary of X and belong to the same minimal parabolic subgroup $G_\xi = Z_\xi N_\xi$, where $Z_\xi = Z_G(A)$, $A = \exp(\mathfrak{a})$ and $N_\xi = N$. Any conjugator from u to v must map $\partial_\infty C$ to itself, hence $g \in G_\xi$ as well. We may therefore write g as $g = a'n$ where $a' \in Z_G(A)$ and $n \in N$.

Since $n \in N$ it follows that we may write n as

$$n = \exp\left(\sum_{\lambda \in \Lambda^+} Z_\lambda\right)$$

where $Z_\lambda \in \mathfrak{g}_\lambda$. When we conjugate u by n we get the following:

$$
\begin{aligned}
nun^{-1} &= \exp\left(\sum_{\lambda \in \Lambda^+} Z_\lambda\right) \exp\left(\sum_{\lambda \in \Pi} Y_\lambda\right) \exp\left(-\sum_{\lambda \in \Lambda^+} Z_\lambda\right) \\
&= \exp\left(\sum_{\lambda \in \Pi} Y_\lambda + \tilde{Y}\right)
\end{aligned}
$$

where $\tilde{Y}$ is the sum of elements $\tilde{Y}_\lambda$ from the non-simple positive root-spaces. Let $\mathfrak{a}$ be the maximal abelian subspace of $\mathfrak{p}$ such that $A = \exp(\mathfrak{a})$. The exponential map, when restricted to $Z_\mathfrak{g}(\mathfrak{a})$, is surjective onto $Z_G(A)$. So there exists $H' \in Z_\mathfrak{g}(\mathfrak{a})$ such that $a' = \exp(H')$. We can decompose $Z_\mathfrak{g}(\mathfrak{a})$ into the direct sum (see, for example, [Ebe96, 2.17.10])

$$Z_\mathfrak{g}(\mathfrak{a}) = \mathfrak{k} \cap Z_\mathfrak{g}(\mathfrak{a}) \oplus \mathfrak{a}.$$

Hence there exists unique $U \in \mathfrak{k} \cap Z_\mathfrak{g}(\mathfrak{a})$ and $H \in \mathfrak{a}$ such that $H' = U + H$. Since U and H commute, $a' = \exp(U)\exp(H) = \exp(H)\exp(U)$. Conjugating nun^{-1} by $\exp(H)$ gives us

$$
\begin{aligned}
\exp(H)nun^{-1}\exp(-H) &= \exp(H)\exp\left(\sum_{\lambda \in \Pi} Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} \tilde{Y}_\lambda\right)\exp(-H) \\
&= \exp\left(\sum_{\lambda \in \Pi} e^{\lambda(H)}Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} e^{\lambda(H)}\tilde{Y}_\lambda\right).
\end{aligned}
$$

Conjugating this by $\exp(U)$ gives us v as

$$v = \exp\left(\sum_{\lambda \in \Pi} e^{\mathrm{ad}(U)}e^{\lambda(H)}Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} e^{\mathrm{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda\right).$$

Notice that, since $U \in Z_\mathfrak{g}(\mathfrak{a})$, for each $\lambda \in \Lambda^+ \setminus \Pi$ the term $e^{\mathrm{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda$ is in the root-space $\mathfrak{g}_\lambda$. But the exponential map gives a bijection between $\mathfrak{n}$ and N. Hence

$$\sum_{\lambda \in \Pi} W_\lambda = \sum_{\lambda \in \Pi} e^{\mathrm{ad}(U)}e^{\lambda(H)}Y_\lambda + \sum_{\lambda \in \Lambda^+ \setminus \Pi} e^{\mathrm{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda.$$

It follows that $W_\lambda = e^{\mathrm{ad}(U)}e^{\lambda(H)}Y_\lambda$ for each simple root λ and $0 = e^{\mathrm{ad}(U)}e^{\lambda(H)}\tilde{Y}_\lambda$ when λ is non-simple. Thus $\tilde{Y} = 0$ and in particular

$$nun^{-1} = u.$$

It follows that

$$v = gug^{-1} = a'nun^{-1}a'^{-1} = a'ua'^{-1}.$$

In order to finish the proof we find an element $H_0 \in \mathfrak{a}$ to do the required job. Let $C_\lambda(U) \in \mathbb{R}$ be such that $[U, Y_\lambda] = C_\lambda(U)Y_\lambda$. Then $e^{\mathrm{ad}(U)}Y_\lambda = e^{C_\lambda(U)}Y_\lambda$ and in particular we see that there exists a positive constant $C_\lambda = e^{C_\lambda(U)+\lambda(H)}$ for each simple root λ such that

$$W_\lambda = C_\lambda Y_\lambda.$$

Now we notice that in $\mathfrak{a}$ we have sufficient degrees of freedom to choose $H_0 \in \mathfrak{a}$ such that $\lambda(H_0) = C_\lambda$ for each $\lambda \in \Pi$. Then H_0 is the required element to complete the proof. $\qquad\square$

Remark: Note that to the existence of the constants $C_\lambda(U)$ required the dimension of each simple root-space in $\mathfrak{g}$ to be equal to 1. So Proposition 8.3 requires $\mathfrak{g}$ to be split.

8.4. **The short conjugators.** Let u, v be unipotent elements contained in the same maximal unipotent subgroup N of G, both of which have all simple entries non-zero. By Algorithm A we can construct g_1 and g_2 in N such that all non-simple entries in $u' = g_1 u g_1^{-1}$ and $v' = g_2 v g_2^{-1}$ are zero. By Proposition 8.3 there exists $g_3 \in A$ such that $g_3 u' g_3^{-1} = v'$. Put $g = g_2^{-1} g_3 g_1$. Then

$$gug^{-1} = v.$$

With this process we can find a short conjugator for u and v.

Theorem 8.4. *Fix $\delta > 0$. Let N be a maximal unipotent subgroup of G. Consider two conjugate unipotent elements*

$$u = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda\right), \quad v = \exp\left(\sum_{\lambda \in \Lambda^+} W_\lambda\right) \in N$$

such that $\min\{\Delta(u), \Delta(v)\} \geq \delta$. Then there exists $g \in G$ such that $gug^{-1} = v$ and which satisfies:

$$d_G(1, g) \leq L(\delta)(d_G(1, u) + d_G(1, v))$$

where $L(\delta)$ will depend on δ and on the root-system Λ associated to G.

Proof. Recall that $g = g_2^{-1} g_3 g_1$ with g_2 and g_1 as in Algorithm A. By Proposition 8.2

$$d_G(1, g_1) + d_G(1, g_2) \leq K(d_G(1, u) + d_G(1, v))$$

where K depends on Λ, c_0 and δ. All we need to do now is obtain a linear upper bound for the size of g_3. By Proposition 8.3 this is member of A, equal to $\exp(H)$ for some $H \in \mathfrak{a}$, which satisfies the following for each simple root λ:

$$(6) \qquad\qquad e^{\lambda(H)} = \frac{\|W_\lambda\|}{\|Y_\lambda\|}$$

where Y_λ is the λ–entry of u and W_λ is the λ–entry of v. The size $d_G(1, g_3)$ is given by the norm of H, which is equal to the Killing form

$$B(H, H) = \text{Trace}(\text{ad}(H)^2) = \sum_{\lambda \in \Lambda} \lambda(H)^2.$$

Since every root in Λ can be expressed as an integer linear combination of simple roots, it follows that there exists a constant S_Λ such that when we take the sum over only the simple roots, rather than all positive roots, we get:

$$(7) \qquad\qquad \sum_{\lambda \in \Pi} \lambda(H)^2 \leq \|H\| = B(H, H) \leq S_\Lambda \sum_{\lambda \in \Pi} \lambda(H)^2.$$

By combining (6) and (7) we get

$$
\begin{aligned}
d_G(1, g_3) &= \|H\| \\
&\leq S_\Lambda \sum_{\lambda \in \Pi} \lambda(H)^2 \\
&= S_\Lambda \sum_{\lambda \in \Pi} (\ln\|W_\lambda\| - \ln\|Y_\lambda\|)^2 \\
&= S_\Lambda \sum_{\lambda \in \Pi} (\ln\|W_\lambda\|)^2 + (\ln\|Y_\lambda\|)^2 - \ln\|W_\lambda\| \ln\|Y_\lambda\| \\
&\leq S_\Lambda \sum_{\lambda \in \Pi} (\ln d_G(1, v))^2 + (\ln d_G(1, u))^2 - 2\ln(\delta)^2
\end{aligned}
$$

This is therefore sufficient to conclude that the size of g, for sufficiently large u, v, is bounded above by a linear function of $d_G(1, u) + d_G(1, v)$, the coefficient of which will depend on δ, $K(\delta)$ and S_Λ. This completes the proof. $\square$

It is well known that the maximal unipotent subgroups in G form one conjugacy class. Furthermore, if we fix a maximal compact subgroup K, then given any pair of maximal unipotent subgroups N_1 and N_2 there exists $k \in K$ such that $kN_1k^{-1} = N_2$. This gives us the following consequence of Theorem 8.4:

Theorem 8.5. *For every $\delta > 0$ there exists a constant $\hat{L}(\delta)$ such that, if u and v are unipotent elements in G satisfying $\min\{\Delta(u), \Delta(v)\} \geq \delta$, then u is conjugate to v if and only if there exists some $g \in G$ such that $gug^{-1} = v$ and*

$$d_G(1, g) \leq \hat{L}(\delta)\big(d_G(1, u) + d_G(1, v)\big).$$

8.5. **Application to lattices.** The condition that $\Delta(u)$ and $\Delta(v)$ must be sufficiently far away from zero is a stronger property than saying they must avoid a neighbourhood of the identity. Nonetheless, with the following Lemma we can use Theorem 8.5 to deduce a result for lattices.

Lemma 8.6. *Let $u \in N$ be as in (3). Then there exists $\delta > 0$ such that if $u \in \Gamma$ then for each simple root $\lambda_i \in \Pi$ either $\|Y_{\lambda_i}\| \geq \delta$ or $Y_{\lambda_i} = 0$.*

Proof. Since $\Gamma \cap N$ is a discrete subgroup of N we know it is finitely generated (see, for example, Corollary 2 of Theorem 2.10 in [Rag72]). Let $\{\gamma_1, \ldots, \gamma_r\}$ be a set of generators for $\Gamma \cap N$ and let $\gamma = \gamma_{i_1}^{\varepsilon_1} \ldots \gamma_{i_s}^{\varepsilon_s} \in \Gamma \cap N$ where $i_j \in \{1, \ldots, r\}$ and $\varepsilon_j \in \mathbb{Z} \setminus \{0\}$. We can write each generator as

$$\gamma_i = \exp\left(\sum_{\lambda \in \Lambda^+} Y_\lambda^{(i)}\right)$$

where $Y_\lambda^{(i)} \in \mathfrak{g}_\lambda$ for each i and each λ. Then, by using the Campbell–Baker–Hausdorff formula,

$$\gamma = \exp\left(\sum_{\lambda \in \Lambda^+} \sum_{j=1}^{s} \varepsilon_j Y_\lambda^{(i_j)} + \tilde{Y}\right)$$

where $\tilde{Y}$ is a sum of terms from non-simple root-spaces. This tells us that each simple entry Y_{λ_i} of u belongs to the integer linear span of the set $\{Y_{\lambda_i}^{(1)}, \ldots, Y_{\lambda_i}^{(r)}\}$, hence there is an element of minimal length for each simple root which can appear as an entry of an element in $\Gamma \cap N$. By taking the shortest of these lengths we obtain a positive value for δ. $\square$

Corollary 8.7. *Let Γ be a lattice in G. Then there exists a constant $L > 0$ such that two unipotent elements u and v in Γ with non-zero simple entries are conjugate in G if and only if there exists a conjugator $g \in G$ such that*

$$d_G(1, g) \leq L\big(d_G(1, u) + d_G(1, v)\big).$$

APPENDIX A. TABLES AND FIGURES FOR LEMMA 7.4
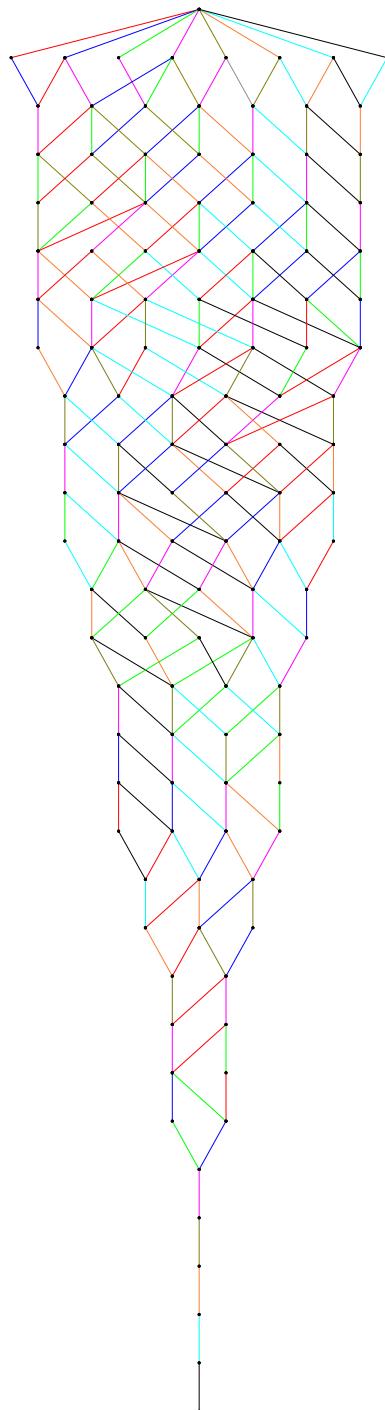
FIGURE 9. A graphical depiction of the positive roots in E_8. The vertices correspond to positive roots (the top vertex is 0), while the edges correspond to addition of a simple root, when reading downwards. Each root has its own colour.

Table 1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
2	1	$\lambda_1 + \lambda_2$	λ_2	
	2	$\lambda_2 + \lambda_4$	λ_4	
	3	$\lambda_3 + \lambda_4$	λ_3	
	4	$\lambda_4 + \lambda_5$	λ_5	
	5	$\lambda_5 + \lambda_6$	λ_6	
	6	$\lambda_6 + \lambda_7$	λ_7	
	7	$\lambda_7 + \lambda_8$	λ_8	
3	4	$\lambda_1 + \lambda_2 + \lambda_4$	$\lambda_1 + \lambda_2$	
	1	$\lambda_2 + \lambda_3 + \lambda_4$	$\lambda_3 + \lambda_4$	
	3	$\lambda_2 + \lambda_4 + \lambda_5$	λ_2	λ_5
	2	$\lambda_3 + \lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2$	
	5	$\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_5 + \lambda_6$	
	6	$\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_6 + \lambda_7$	
	7	$\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_7 + \lambda_8$	
4	1	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	$\lambda_1 + \lambda_2 + \lambda_4$	
	3	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2$	λ_4
	2	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	$\lambda_2 + \lambda_4 + \lambda_5$	
	5	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_2 + \lambda_4$	λ_6
	4	$\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_4 + \lambda_5 + \lambda_6$	
	6	$\lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_5 + \lambda_6 + \lambda_7$	
	7	$\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_6 + \lambda_7 + \lambda_8$	
5	3	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	
	1	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	
	4	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + \lambda_2$	$\lambda_5 + \lambda_6$
	2	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6$	
	6	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_4$	$\lambda_6 + \lambda_7$
	5	$\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	
	7	$\lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
6	3	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$	
	5	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + \lambda_2 + \lambda_3$	$\lambda_5 + \lambda_6$
	2	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	
	4	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5 + \lambda_6$	
	1	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7$	
	7	$\lambda_2 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_4 + \lambda_5$	$\lambda_7 + \lambda_8$

Table 1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	6	$\lambda_3 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_4 + \lambda_5 + \lambda_6$ $+\lambda_7 + \lambda_8$	
7	1	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5$	$\lambda_1 + \lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5$	
	3	$\lambda_1 + \lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + \lambda_2 + \lambda_3$ $+\lambda_4 + \lambda_5 + \lambda_6$	
	2	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6$	
	4	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + \lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	
	6	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$	λ_7
	5	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	
	7	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_3 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	
8	5	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5$	
	1	$\lambda_1 + \lambda_2 + \lambda_3$ $+2\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6$	
	3	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	
	2	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	
	6	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	
9	4	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5 + \lambda_6$	
	5	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + \lambda_5$	λ_7
	3	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	
	6	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	1	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7$	$\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	
	2	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + \lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	
10	1	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+2\lambda_4 + 2\lambda_5 + \lambda_6$	
	2	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	

Table 1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	6	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7$	λ_6
	4	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+\lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7$	
11	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	$\lambda_1 + 2\lambda_2 + \lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	
	2	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	
	6	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + \lambda_7 + \lambda_8$	
12	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6$	
	2	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5$ $+2\lambda_6 + 2\lambda_7 + \lambda_8$	
13	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + \lambda_6 + \lambda_7$	
	3	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	1	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7$	
	4	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	5	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + \lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
14	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 2\lambda_5 + 2\lambda_6 + \lambda_7$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 2\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	

Table 1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
15	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+4\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+3\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8$	λ_7
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
16	1	$\lambda_1 + 3\lambda_2 + 2\lambda_3$ $+4\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$ $+4\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+2\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
	3	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
17	1	$2\lambda_1 + 3\lambda_2 + 2\lambda_3$ $+4\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	$\lambda_1 + 3\lambda_2 + 2\lambda_3$ $+4\lambda_4 + 3\lambda_5 + 2\lambda_6 + \lambda_7$	
	2	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	3	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
	4	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + \lambda_3 + 3\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
18	1	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	2	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
	3	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
19	3	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + \lambda_7 + \lambda_8$	
	2	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 2\lambda_6 + 2\lambda_7 + \lambda_8$	
	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
20	1	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
	2	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
21	2	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+3\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
	1	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
22	2	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 4\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	

Table 1: The simple case for root systems of type E_8

Height	Order	λ	μ or μ_1	μ_2 (if needed)
	1	$\lambda_1 + 3\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
23	1	$2\lambda_1 + 4\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 3\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
	2	$2\lambda_1 + 3\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$\lambda_1 + 3\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
24	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 2\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
25	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 5\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
26	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+4\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
27	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 2\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 3\lambda_6 + 2\lambda_7 + \lambda_8$	
28	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 3\lambda_7 + \lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 2\lambda_7 + \lambda_8$	
29	1	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 3\lambda_7 + 2\lambda_8$	$2\lambda_1 + 4\lambda_2 + 3\lambda_3 + 6\lambda_4$ $+5\lambda_5 + 4\lambda_6 + 3\lambda_7 + \lambda_8$	

TABLE 2. The simple case for root systems of type F_4

| Height	Order	λ	μ or μ_1	μ_2 (if needed)
2	1	$\lambda_1 + \lambda_2$	λ_2	
	2	$\lambda_2 + \lambda_3$	λ_3	
	3	$\lambda_3 + \lambda_4$	λ_4	
3	2	$\lambda_1 + \lambda_2 + \lambda_3$	$\lambda_1 + \lambda_2$	
	1	$\lambda_2 + 2\lambda_3$	$\lambda_2 + \lambda_3$	
	3	$\lambda_2 + \lambda_3 + \lambda_4$	$\lambda_3 + \lambda_4$	
4	1	$\lambda_1 + \lambda_2 + 2\lambda_3$	$\lambda_1 + \lambda_2 + \lambda_3$	
	3	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	$\lambda_1 + \lambda_2$	λ_4
	2	$\lambda_2 + 2\lambda_3 + \lambda_4$	$\lambda_2 + \lambda_3 + \lambda_4$	
5	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3$	$\lambda_1 + \lambda_2 + 2\lambda_3$	
	3	$\lambda_1 + \lambda_2 + 2\lambda_3 + \lambda_4$	$\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$	
	2	$\lambda_2 + 2\lambda_3 + 2\lambda_4$	$\lambda_2 + 2\lambda_3 + \lambda_4$	
6	1	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + \lambda_4$	$\lambda_1 + 2\lambda_2 + 2\lambda_3$	
	2	$\lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4$	$\lambda_2 + 2\lambda_3 + 2\lambda_4$	
7	1	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + \lambda_4$	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + \lambda_4$	
	2	$\lambda_1 + 2\lambda_2 + 2\lambda_3 + 2\lambda_4$	$\lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4$	
8	1	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + 2\lambda_4$	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + \lambda_4$	
9	1	$\lambda_1 + 3\lambda_2 + 3\lambda_3 + 2\lambda_4$	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + 2\lambda_4$	
10	1	$\lambda_1 + 3\lambda_2 + 4\lambda_3 + 2\lambda_4$	$\lambda_1 + 2\lambda_2 + 3\lambda_3 + 2\lambda_4$	
11	1	$2\lambda_1 + 3\lambda_2 + 4\lambda_3 + 2\lambda_4$	$\lambda_1 + 3\lambda_2 + 4\lambda_3 + 2\lambda_4$	
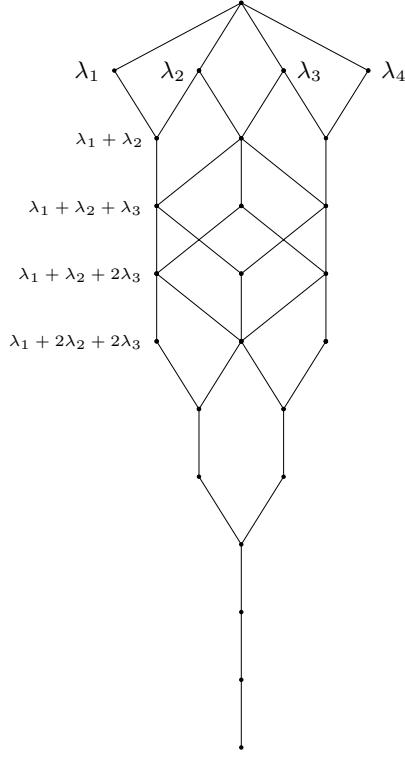
FIGURE 10. A graphical depiction of the positive roots in F_4. The horizontal levels correspond to heights. The elements of height one are labelled, and one more in each height up to 5, but the rest are not. When you move down a height, following an edge corresponds to adding a simple root.
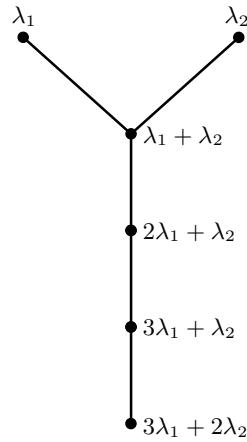

FIGURE 11. A graphical depiction of the positive roots in G_2.

REFERENCES

[BD11]   Jason Behrstock and Cornelia Drutu. Divergence, thick groups, and short conjugators. *arXiv:1110.5005v1 [math.GT]*, 2011.

[BGS85] Werner Ballmann, Mikhael Gromov, and Viktor Schroeder. *Manifolds of nonpositive curvature*, volume 61 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1985.

[BH99]   Martin R. Bridson and André Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[CGW09]   John Crisp, Eddy Godelle, and Bert Wiest. The conjugacy problem in subgroups of right-angled artin groups. *Journal of Topology*, 2(3):442–460, 2009.

[Ebe96]   Patrick B. Eberlein. *Geometry of nonpositively curved manifolds*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1996.

[GI05]   F. Grunewald and N. Iyudu. The conjugacy problem for two-by-two matrices over polynomial rings. *Sovrem. Mat. Prilozh.*, (30, Algebra):31–45, 2005.

[GS80]   Fritz Grunewald and Daniel Segal. Some general algorithms. I. Arithmetic groups. *Ann. of Math. (2)*, 112(3):531–583, 1980.

[Hel01]   Sigurdur Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 34 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Corrected reprint of the 1978 original.

[HKM10]   Thomas J. Haines, Michael Kapovich, and John J. Millson. Ideal triangles in Euclidean buildings and branching to Levi subgroups. *arXiv:1011.6636v1 [math.RT]*, 2010.

[JOR10]   Ronghui Ji, Crichton Ogle, and Bobby Ramsey. Relatively hyperbolic groups, rapid decay algebras and a generalization of the Bass conjecture. *J. Noncommut. Geom.*, 4(1):83–124, 2010. With an appendix by Ogle.

[KL97]   Bruce Kleiner and Bernhard Leeb. Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings. *Inst. Hautes Études Sci. Publ. Math.*, (86):115–197 (1998), 1997.

[KT04]   Linus Kramer and Katrin Tent. Asymptotic cones and ultrapowers of Lie groups. *Bull. Symbolic Logic*, 10(2):175–185, 2004.

[Leu04]   Enrico Leuzinger. Tits geometry, arithmetic groups, and the proof of a conjecture of Siegel. *J. Lie Theory*, 14(2):317–338, 2004.

[LMR00]   Alexander Lubotzky, Shahar Mozes, and M. S. Raghunathan. The word and Riemannian metrics on lattices of semisimple groups. *Inst. Hautes Études Sci. Publ. Math.*, (91):5–53 (2001), 2000.

[Lus07]   Martin Lustig. Conjugacy and centralizers for iwip automorphisms of free groups. In *Geometric Group Theory*, Trends in Mathematics, pages 197–224. Birkhäuser Basel, 2007.

[MM00]   H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.

[Par00]   Anne Parreau. *Dégénérescences de sous-groupes discrets de groupes de Lie semisimples et actions de groupes sur des immeubles affines*. PhD thesis, Univ. Orsay, 2000.

[Rag72]   M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, New York, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68.

[Sal12a]   Andrew Sale. Conjugacy length in group extensions. 2012.

[Sal12b]   Andrew Sale. *The Length of Conjugators in Solvable Groups and Lattices of Semisimple Lie Groups*. DPhil thesis, University of Oxford, 2012.

[Sal12c]   Andrew Sale. On the Magnus embedding and the conjugacy length function of wreath products and free solvable groups. *arXiv:1202.5343v2*, 2012.

[Tao11]   Jing Tao. Linearly bounded conjugator property for mapping class groups. *arXiv:1106.2341v2 [math.GT]*, 2011.

E-mail address: `andrew.sale@some.oxon.org`